# UNIVERSIDAD COMPLUTENSE DE MADRID

FACULTAD DE MEDICINA

Departamento de Microbiología I


## TESIS DOCTORAL

# Enfermedad inflamatoria intestinal y celiaquía : componente genético y farmacogenética

MEMORIA PARA OPTAR AL GRADO DE DOCTOR

PRESENTADA POR

**Luz María Medrano de Dios**

Directoras

Elena Urcelay García
Concepción Núñez Pardo de Vera

**Madrid, 2014**

**UNIVERSIDAD COMPLUTENSE DE MADRID**

**FACULTAD DE MEDICINA**
**DEPARTAMENTO DE MICROBIOLOGÍA I**

# Enfermedad inflamatoria intestinal y celiaquía: componente genético y farmacogenética

Memoria presentada por

**Luz María Medrano de Dios**

para optar al grado de Doctor

DIRECTORES

Elena Urcelay García

Concepción Núñez Pardo de Vera

Servicio de Inmunología (Hospital Clínico San Carlos)

Madrid, 2014

La realización de este trabajo se ha visto facilitada gracias a la financiación otorgada por la Fundación Eugenio Rodríguez Pascual (Estudios genéticos en pacientes de enfermedad celíaca. Comparación con enfermedad inflamatoria intestinal) y la Fundación Mutua Madrileña (Farmacogenética del tratamiento con infliximab en pacientes de Crohn)

# Informe del Director de la Tesis Doctoral

| *DATOS DE LA TESIS DOCTORAL* | |
|---|---|
| *Nombre del Doctorando* | **LUZ MARÍA MEDRANO DE DIOS** |
| *Título de la Tesis* | **ENFERMEDAD INFLAMATORIA INTESTINAL Y CELIAQUÍA: COMPONENTE GENÉTICO Y FARMACOGENÉTICA.** |
| *Facultad o Centro* | **FACULTAD DE MEDICINA (UCM)** |

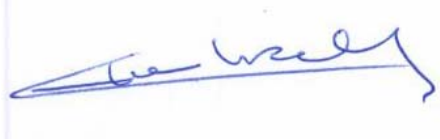
| *DATOS DEL DIRECTOR DE LA TESIS DOCTORAL* | |
|---|---|
| *Nombre Completo* | **ELENA URCELAY GARCÍA** |
| *Centro al que pertenece y dirección* | **Instituto de Investigación Sanitaria del Hospital Clínico San Carlos (IdISSC)**<br>**Hospital Clínico San Carlos, Servicio de Inmunología**<br>**c/ Profesor Martín Lagos s/n**<br>**28040 Madrid** |
| *D.N.I/Pasaporte* | *07219750G* |
| *e-mail* | **elena.urcelay@salud.madrid.org** |

| | *VALORACIÓN DE LA TESIS* | | | |
|---|---|---|---|---|
| | *Muy Buena* | *Buena* | *Suficiente* | *Deficiente* |
| *Originalidad* | | *x* | | |
| *Definición Objetivos* | *x* | | | |
| *Metodología* | *x* | | | |
| *Relevancia Resultados* | *x* | | | |
| *Discusión / Conclusiones* | *x* | | | |

***INFORME*** *(en caso necesario se podrán añadir más hojas):*

***Los objetivos de esta tesis son realistas y está planteada y desarrollada con rigor. La metodología es adecuada al volumen de enfermos estudiados. Los resultados obtenidos contribuirán sin duda a mejorar el conocimiento de las enfermedades y, como consecuencia, a incrementar la calidad de vida de los afectados. Las conclusiones dan respuesta al planteamiento inicial y la bibliografía está bien actualizada.***
***Es, por tanto, un trabajo de factura impecable con una presentación clara a pesar de la dificultad inherente a los datos.***

***Madrid, a 2 de Diciembre de 2013***

***Fdo.:*** 

***Este impreso deberá entregarse al Departamento/Órgano responsable del Posgrado/ Comisión responsable del Programa de Doctorado, para su estudio y aprobación en la admisión a trámite de la tesis doctoral. Asimismo, deberá incluirse entre la documentación enviada a la Comisión de Doctorado para la designación del Tribunal y aprobación de la defensa de la Tesis Doctoral.***

# Informe del Director de la Tesis Doctoral

| *DATOS DE LA TESIS DOCTORAL* | |
|---|---|
| *Nombre del Doctorando* | **LUZ MARÍA MEDRANO DE DIOS** |
| *Título de la Tesis* | **ENFERMEDAD INFLAMATORIA INTESTINAL Y CELIAQUÍA: COMPONENTE GENÉTICO Y FARMACOGENÉTICA.** |
| *Facultad o Centro* | **FACULTAD DE MEDICINA (UCM)** |

| *DATOS DEL DIRECTOR DE LA TESIS DOCTORAL* | |
|---|---|
| *Nombre Completo* | **CONCEPCIÓN NÚÑEZ PARDO DE VERA** |
| *Centro al que pertenece y dirección* | **Instituto de Investigación Sanitaria San Carlos (IdISSC)**<br>**Hospital Clínico San Carlos, servicio de inmunología, c/ profesor martín lagos s/n 28040 Madrid** |
| *D.N.I/Pasaporte* | ***33299125R*** |
| *e-mail* | ***conchita.npardo@gmail.com*** |

| | *VALORACIÓN DE LA TESIS* | | | |
|---|---|---|---|---|
| | *Muy Buena* | *Buena* | *Suficiente* | *Deficiente* |
| *Originalidad* | | *X* | | |
| *Definición Objetivos* | *X* | | | |
| *Metodología* | | *X* | | |
| *Relevancia Resultados* | | *X* | | |
| *Discusión / Conclusiones* | | *X* | | |

***INFORME*** *(en caso necesario se podrán añadir más hojas):*

La Dra. M. Concepción Núñez Pardo de Vera, la cual trabaja en el Servicio de Inmunología Clínica del Hospital Clínico San Carlos

INFORMA: que Luz María Medrano de Dios, licenciada en Ciencias Biológicas por la Universidad de Salamanca, ha realizado bajo mi dirección y la de la Dra. Elena Urcelay el trabajo de investigación titulado: "Enfermedad inflamatoria intestinal y celiaquía: componente genético y farmacogenética", que considero cumple los requisitos necesarios para ser defendido como tesis doctoral.

***Madrid, a 2 de Diciembre de 2.013***

***Fdo.:*** 

***Este impreso deberá entregarse al Departamento/Órgano responsable del Posgrado/ Comisión responsable del Programa de Doctorado, para su estudio y aprobación en la admisión a trámite de la tesis doctoral. Asimismo, deberá incluirse entre la documentación enviada a la Comisión de Doctorado para la designación del Tribunal y aprobación de la defensa de la Tesis Doctoral.***

*Dedicada a mis padres…*

*Agradecimientos*

Quiero dar las gracias en primer lugar a mis directoras de tesis la Dra. Urcelay y Dra. Núñez por la oportunidad que me han brindado en participar en sus líneas de investigación, sin sus enseñanzas este trabajo no hubiera sido posible.

Quiero hacer un agradecimiento especial a Conchi, con la persona que he compartido tanto tiempo en este trabajo, por tener tanta paciencia en las explicaciones, por tener siempre un minuto de su tiempo, por escucharme, por ser mi apoyo en las decisiones finales, por comprenderme en mi recta final, en definitiva por ser mi amiga, gracias.

El camino ha sido largo y no hubiese sido posible sin mis compañeros a los que tengo que agradecer tanto: Manmen esa chica granadina que volví a encontrarme en Holanda y tantas anécdotas nos han pasado, que sería de nosotras sin el apoyo mutuo en los *monday meetings*. Carlos, gracias por hacerme sentir como en mi casa al llegar a este laboratorio. Laura, gracias por haber estado ahí para resolver cualquier tipo de duda y haber sido tan concisa. Arturo, uno de los mejores inmunólogos que he conocido y al que tanto se ha echado de menos en su regreso a casa. Gorka, la persona que siempre está dispuesta a explicar sus amplios conocimientos. Romina, mi co-R y con la que he tenido que resolver situaciones difíciles en los cuatro años de residencia. Miguel Ángel (Miki para los amigos), no tengo palabras para expresar lo mucho que estoy agradecida en todo el apoyo que me has dado desde que nos conocimos, continuaremos hablando del futuro y conseguiremos todo lo que nos propongamos. María, única en el servicio de Inmunología. Lidia, Vir, Belén, Ale y Juli, por las tardes en el laboratorio pensando en el dulce que elegir, por las sobremesas, por los momentos de risas en la biblioteca, en definitiva un agradecimiento "VIP" a todas esas personas que entre escritura y escritura me hicieron reír. También quiero agradecer a Ale y Juli por haber tenido siempre un momento para escucharme en mi recta final, por nuestras cañitas los fines de semana y por haber inundado de alegría el servicio de Inmunología, gracias. Félix, siempre agradecida con tus visitas.

También quiero agradecer a todo el personal del servicio de Inmunología, sin ellos mucho del trabajo que presento no podría haber sido posible: Emilio, María Ángeles, José Luis Subiza, Miguel, Silvia, Amparo, Ángeles, Menchu, Carmen Poyo, Pilar, Toñi, Pepa, Carmen Martínez, Nieves, Bárbara, María, Alejandro, Ángel, José Luis Santiago, Jezabel y Priscila.

I would like to thank all people who I met in Groningen and helped me during my three months there. Thanks to Dr Wijmenga for the opportunity of working in her laboratory and to my colleagues: Isis, Javier, Manmen, Thays, Rodrigo, Cleo, Anna and Juha for the moments in the “Neighbour” and in the president´s room. Thank you.

No quería terminar sin agradecer a las personas más importantes en mi vida, las personas que me han hecho hoy lo que soy, mis padres y hermana, gracias a ellos por apoyarme en todas mis decisiones, por darme la fuerza necesaria para terminar todos los proyectos en mi vida. Os quiero.

*Índice*

| | |
|---|---|
| APC | Células presentadoras de antígeno *(antigen presenting cells)* |
| AR | Artritis reumatoide |
| *C1orf106* | *Chromosome 1 open reading frame 106* |
| *CARD15* | *Caspase recruitment domain-containing protein 15* |
| *CARD9* | *Caspase recruitment domain family, member 9* |
| *CCR6* | *C-C chemokine receptor type 6* |
| *CDKN2B* | *Cyclin-dependent kinase inhibitor 2B (p15, inhibits CDK4)* |
| *CTLA4* | *Cytotoxic T-lymphocyte-associated protein 4* |
| CU | Colitis ulcerosa |
| *CUL2* | *Cullin 2* |
| CNV | *Copy number variation* |
| DM1 | Diabetes tipo 1 |
| DM2 | Diabetes tipo 2 |
| EAE | Encefalomielitis autoinmune experimental |
| EATL | Enteropatía asociada a linfoma T *(Enteropathy associated T cell lymphoma)* |
| EC | Enfermedad de Crohn |
| ECe | Enfermedad celiaca |
| ECeR | Enfermedad celiaca refractaria |
| EII | Enfermedad inflamatoria intestinal |
| EM | Esclerosis múltiple |
| ENCODE | *Encyclopedia of DNA Elements* |
| eQTL | *Expression quantitative trait loci* |
| *ERAP1* | *Endoplasmic reticulum aminopeptidase 1* |
| ESPGHAN | *European Society for Paediatric Gastroenterology, Hepatology and Nutrition* |
| *G0S2* | *G0/G1switch 2* |
| *GLB1* | *Galactosidase, beta 1* |
| GWAS | Estudios de asociación de barrido genómico *(Genome Wide Association Studies)* |
| HBD | b-defensina humanas |
| HLA | Antígeno Leucocitario Humano *(Human Leukocyte Antigen)* |
| *HNF4A* | *Hepatocyte nuclear factor 4, alpha* |
| *ICOS* | *Inducible T-cell co-stimulator* |
| IFN | Interferón |
| IFN-γ | Interferón gamma |
| IFX | Infliximab |
| Ig | Inmunoglobulina |
| IL- | Interleuquina (*interleukin*) |
| *IL12RB1* | *Interleukin 12 receptor, beta 1* |
| *IL17RA* | *Interleukin 17 receptor A* |
| *IL18RAP* | *Interleukin 18 receptor accessory protein* |
| *IL23R* | *Interleukin 23 receptorIL2RA Interleukin 2 receptor, alpha* |
| *IL6R* | *Interleukin 6 receptor* |

| | |
|---|---|
| *IRF1* | *Interferon regulatory factor 1* |
| *JAK2* | *Janus kinase 2* |
| LD | Desequilibrio de ligamiento *(linkage desequilibrium)* |
| LIEs | Linfocitos intraepiteliales |
| lincRNA | *Large intergenic non-coding RNAs* |
| *LPP* | *LIM domain containing preferred translocation partner in lipoma* |
| MDR | Multifactor dimensionality reduction |
| MHC | Complejo Principal de Histocompatibilidad *(Major Histocompatibility Complex)* |
| miRNA | *micro RNA* |
| MMPs | Metaloproteasas de la matriz |
| *MUC19* | *Mucin 19, oligomeric* |
| NADPH | Nicotinamida adenina dinucleótido fosfato *(Nicotinamide adenine dinucleotide phosphate)* |
| NK | Células citotóxicas naturales *(natural killer)* |
| *NLR* | *NOD-like receptors* |
| *NOD2* | *Nucleotide-binding oligomerization domain containing 2* |
| *OCTN1, OCTN2* | *Organic cation transporter 1,2* |
| PBMC | Célula mononuclear de sangre periférica *(Peripheral blood mononuclear cell)* |
| *PTPN22* | *Protein tyrosine phosphatase, non-receptor type 22 (lymphoid)* |
| *PUS10* | *Pseudouridylate synthase 10* |
| *RHBDD3* | *Rhomboid domain containing 3* |
| *RNF186* | *Ring finger protein 186* |
| *RORC* | *RAR-related orphan receptor C* |
| *ROR-γt* | *Receptor tyrosine kinase-like orphan receptor gamma t* |
| *S100A8* | *S100 calcium binding protein A8* |
| *S100A9* | *S100 calcium binding protein A9* |
| *SCHIP1* | *Schwannomin interacting protein 1* |
| *SMAD3* | *SMAD family member 3* |
| SNP | Polimorfismo de un único nucleótido *(single nucleotide polymorphism)* |
| *SOCS3* | *Suppressor of cytokine signaling 3* |
| *STAT3* | *Signal transducer and activator of transcription 3* |
| *TAGAP* | *T-cell activation RhoGTPase activating protein* |
| *TBX21* | *T-box 21* |
| *TGFb* | *Transforming growth factor beta* |
| Th | Linfocito T colaborador *(Thelper)* |
| *TLR* | *Toll-like-receptors* |
| *TNF-a* | *Tumor necrosis factor a* |
| *TNFAIP6* | *Tumor necrosis factor, alpha-induced protein 6* |
| *TNFRSF1A* | *Tumor necrosis factor receptor superfamily, member 1A* |
| *TNFRSF1B* | *Tumor necrosis factor receptor superfamily, member 1B* |
| *TNFSF15* | *Tumor necrosis factor (ligand) superfamily, member 15* |

| | |
|---|---|
| Treg | Linfocito T regulador |
| *TYK2* | *Tyrosine kinase 2* |
| WES | Secuenciación del exoma completo *(Whole exome sequencing)* |
| WGS | Secuenciación del genoma completo *(Whole genome sequencing)* |
| *WTCCC* | *Welcome Trust Case-Control Consortium* |

***Resumen/Summary***

**Introducción**

La enfermedad inflamatoria intestinal (EII) y la enfermedad celiaca (ECe) son dos enfermedades mediadas por el sistema inmunológico con afectación intestinal crónica. En el caso de la ECe es conocido el papel etiológico del gluten en el desarrollo de la enfermedad, mientras que en la EII una de las causas principales parece ser una alteración en la respuesta inmune frente a flora intestinal.

En los últimos años los avances en los estudios genéticos, gracias principalmente al desarrollo de estudios a gran escala (GWAS), han permitido identificar un amplio número de variantes asociadas a estas enfermedades, hasta 163 loci en EII y 41 en ECe. A pesar de ello, existe un porcentaje de heredabilidad aún por descubrir ("heredabilidad perdida") en el cual pueden estar involucrados diferentes factores, como son las interacciones epistáticas entre genes o con factores ambientales. Además, puede haber factores genéticos que afecten a grupos específicos de pacientes, situación que no se ha contemplado en los GWAS realizados.

En el caso de la ECe, el tratamiento consiste en la eliminación de gluten de la dieta, que resulta eficaz en casi la totalidad de los individuos. Sin embargo, para EII existen diversos tratamientos que presentan distinto grado de eficacia. En concreto, en pacientes con enfermedad de Crohn (EC) uno de los fármacos administrados cuando fallan los tratamientos convencionales es el infliximab (IFX) (anticuerpo monoclonal anti-TNF-$\alpha$), pero la respuesta a este fármaco varía dependiendo del tiempo de evolución de la enfermedad y de otros factores, muy probablemente factores genéticos entre ellos.

**Hipótesis**

- Las interacciones epistáticas entre genes pueden ser responsables en parte del componente genético sin explicar, denominado "heredabilidad perdida"; en enfermedades multifactoriales como la ECe o la EII.
- Alteraciones de la respuesta inmune Th17 pueden estar involucradas en el origen de la ECe. Los estudios previos de GWAS no parecen apoyar esta idea, pero ello no excluye su implicación debido a las posibles interacciones entre genes involucrados en la ruta Th17 no analizadas en los estudios de GWAS.

- El estudio de la expresión a nivel intestinal de genes asociados a ECe o EII puede ayudar a comprender mejor cómo se desencadenan o desarrollan ambas enfermedades.
- El perfil genético de cada individuo puede influir en el tipo de respuesta a IFX.

**Objetivos**

✓ Búsqueda de factores genéticos de susceptibilidad a enfermedad celiaca o enfermedad inflamatoria intestinal que ayuden a explicar la "heredabilidad perdida":

- Búsqueda de nuevos factores de susceptibilidad a enfermedad celiaca en la región del HLA.
- Confirmación de interacciones gen-gen (*IL23R*, *CARD15* y *TLR9*) como factores de riesgo a enfermedad inflamatoria intestinal.
- Estudio de la asociación de genes que intervienen en la ruta Th17 con la susceptibilidad a padecer enfermedad celiaca.

✓ Estudio de la expresión a nivel intestinal de genes implicados en la ruta Th17 en enfermedad celiaca y en colitis ulcerosa.

✓ Estudio de factores genéticos que ayuden a predecir la respuesta a infliximab en pacientes con enfermedad de Crohn.

**Material y Métodos**

1. <u>Búsqueda de nuevos factores de susceptibilidad a enfermedad celiaca en la región del HLA.</u>

   Estudiamos 274 trios (en los que se incluye el individuo afectado y sus dos progenitores), 369 pacientes con ECe y 461 individuos sin enfermedades mediadas por el sistema inmunológico que utilizamos como controles. Se genotiparon en todas las muestras los genes *HLA-DRB1*, *-DQA1*, *-DQB1, -B8,* los microsatélites *TNFa* y *TNFb* y los polimorfismos del *TNF* -308 (rs1800629) y -376 (rs1800750) para identificar la presencia de los dos principales haplotipos que contienen *DRB1*03:01* en los individuos estudiados. Los datos de las familias se analizaron mediante el *transmission disequilibrium test* (TDT)

y los datos caso-control mediante el test chi-cuadrado o el test exacto de Fisher si algún valor esperado era menor de 5.

2. Confirmación de interacciones gen-gen (*IL23R*, *CARD15* y *TLR9*) como factores de riesgo a enfermedad inflamatoria intestinal.
   Estudiamos una muestra inicial de 416 pacientes de EC, 452 pacientes con CU y 547 individuos sin enfermedades mediadas por el sistema inmunológico que utilizamos como controles y una segunda muestra consistente en 122 pacientes de EC y 307 controles. Por medio de tecnología TaqMan se genotiparon 9 polimorfismos de un único nucleótido (SNPs) presentes en los genes *CARD15*, *IL23R* y *TLR9.* Los estudios estadísticos de asociación fueron realizados con el test chi-cuadrado o test exacto de Fisher.

3. Estudio de la asociación de genes que intervienen en la ruta Th17 con la susceptibilidad a padecer enfermedad celiaca.
   Para este estudio empleamos una primera muestra de 735 individuos con ECe y 549 controles y un segundo grupo de 294 individuos con ECe y 475 controles. Se llevó a cabo el análisis de 101 polimorfismos presentes en 16 genes (*CCR6*, *GLB1*, *IL17F*, *IL2*, *IL21*, *IL22*, *IL23A*, *IL23R*, *IL6*, *JAK2*, *RHBDD3*, *SMAD3*, *STAT3* y *TYK2*) por medio de tecnología Veracode. Los SNPs se seleccionaron tratando de recoger la máxima variabilidad en todas las regiones por medio de un *aggressive tagging* en todos los genes con excepción de *IL6R, JAK2* y *STAT3.* El estudio de asociación caso-control fue analizado por medio de test chi-cuadrado o el test exacto de Fisher y el análisis de interacciones gen-gen se llevó a cabo por regresión logística, *random forests* (RF), *classification and regression trees* (CART) y *multifactor dimensionality reduction* (MDR).

4. Estudio de la expresión a nivel intestinal de genes implicados en la ruta Th17 en enfermedad celiaca y en colitis ulcerosa.
   Medimos los niveles de expresión de 12 genes relacionados con la ruta Th17 (*CCR6*, *IL17F*, *IL2*, *IL21*, *IL22*, *IL23A*, *IL23R*, *IL6*, *JAK2*, *RHBDD3*, *SMAD3* y *STAT3)* en 24 biopsias duodenales de pacientes con ECe y 6 de individuos sanos por medio de PCR cuantitativa usando tecnología TaqMan. A su vez estratificamos las biopsias de ECe dependiendo de la edad de debut del paciente

(14 de pacientes pediátricos y 6 de adultos) y del tratamiento (4 de pacientes en dieta sin gluten y 20 en dieta con gluten). Los genes *GLB1* y *TYK2* se utilizaron como genes de referencia para la normalización de nuestros resultados tras el análisis de 48 genes mediante los algoritmos geNorm versión 3.5 y NormFinder versión 0.953 presentes en el software Genex. Los niveles de expresión se determinaron mediante el método ΔΔCt ($\Delta Ct_{gen\ target}$-$\Delta Ct_{gen\ referencia}$). El estudio comparativo de las muestras se llevó a cabo con el test U de Mann Whitney utilizando el programa SPSS 15.1.

Los niveles de expresión de los 12 genes mencionados se analizaron también en muestras de colon o recto procedentes de 8 individuos con CU. De cada paciente se analizó una muestra de tejido normal y una de tejido inflamado. El análisis estadístico se llevó a cabo por medio del test de Wilcoxon (de muestras pareadas) presente en el paquete estadístico SPSS.

5. Estudios de farmacogenética en pacientes con EC tratados con IFX.

Un primer trabajo de farmacogenética se llevó a cabo incluyendo 238 pacientes de EC respondedores a IFX y 59 no respondedores. A todos los pacientes se les administró IFX a las 0, 2 y 6 semanas y la respuesta al tratamiento se evaluó a la semana 10 tras la primera dosis de IFX por medio del índice de Harvey-Bradshaw (IHB). La respuesta parcial se consideró con valores de IHB<3 y la remisión con valores de IHB≤4. Se analizaron 4 SNPs presentes en el gen *TNFRSF1B* (rs1061622, rs1061624 y rs3397) y *TNFRSF1A* (rs767455) por medio de tecnología TaqMan (Applied Biosystems, Foster City, CA, USA). Los haplotipos entre los SNPs de *TNFRSF1B* en desequilibrio de ligamiento se determinaron mediante el algoritmo de Expectación-Maximación utilizando el programa Haploview 4.1. La comparación de las frecuencias alélicas y haplotípicas entre respondedores y no respondedores se llevó a cabo por tests chi-cuadrado o el test exacto de Fisher. Características demográficas de los pacientes como fumar, clasificación de Montreal, años de evolución, tratamiento concomitante… se evaluaron como posibles variables de confusión por medio de regresión logística binaria utilizando el programa estadístico SPSS15.1 .

El último estudio de farmacogenética incluyó el estudio de los genes *TNFAIP6*, *IL11*, *G0S2*, *S100A8* y *S100A9*. Para ello contamos con un total de 210 pacientes tratados con IFX (166 respondedores y 44 no respondedores), las pautas de

administración, respuesta al fármaco y estudio estadístico fueron las mismas que el estudio anterior.

**Resultados y discusión**

1. Búsqueda de nuevos factores de susceptibilidad a enfermedad celiaca en la región del HLA.

   Observamos que los individuos que presentan el haplotipo ancestral (AH) 8.1 presentan mayor riesgo de desarrollar ECe que individuos que portan la molécula DQ2.5 pero con un haplotipo diferente. Por tanto parece que existe un factor de riesgo adicional en el AH 8.1. El efecto de este factor sólo se manifiesta en individuos que no presentan una segunda copia del alelo *DQB1*02*. Este factor podría estar también presente en los individuos DQ2.5 *trans* debido a que ambos grupos presentan el mismo riesgo a desarrollar ECe. Con la utilización de los datos del proyecto Inmunochip tratamos de buscar esa posible variante compartida entre los individuos AH8.1 y DQ2.5 *trans*, pero no pudimos determinar tal variante probablemente debido al insuficiente tamaño muestral dada la complejidad de la región HLA. Estos resultados se han recogido en el artículo: *HLA and celiac disease susceptibility: new genetic factors bring open questions about the HLA influence and gene-dosage effects* PLoS One. 2012;7(10):e48403.

2. Confirmación de interacciones gen-gen (*IL23R*, *CARD15* y *TLR9*) como factores de riesgo a enfermedad inflamatoria intestinal.

   Nuestros resultados muestran que el alelo minoritario del SNP rs7517847 del gen *IL23R* disminuye el riesgo a desarrollar EC en aquellos individuos homozigotos rs352162_CC de *TLR9* (p=0,0003 OR=0,44 95% CI 0,27-0,71). Por tanto observamos una interacción epistática entre los genes *IL23R* y *TLR9* asociada al riesgo a EC, similar a lo observado en población alemana. Por el contrario no vemos interacción entre variantes presentes en *TLR9* y *CARD15*. Estos resultados son de gran interés debido a la implicación funcional de ambos genes en el reconocimiento y defensa de patógenos. El manuscrito que recoge estos datos se está enviando para su publicación.

3. Estudio de la asociación de genes que intervienen en la ruta Th17 con la susceptibilidad a padecer enfermedad celiaca.

   No observamos asociación de ninguno de los 101 SNPs analizados en los 16 genes implicados en la respuesta Th17 ni por sí solos ni interaccionando con otros genes de la misma ruta. Con lo que podemos confirmar que no existe una implicación genética de la ruta Th17 en el origen de la ECe, así se recoge en el artículo publicado: *Th17-related genes and celiac disease susceptibility.* PLoS One. 2012;7(2):e31244.

4. Estudio de la expresión a nivel intestinal de genes implicados en la ruta Th17 en ECe y en CU.

   Observamos una expresión diferente de los genes *IL23A*, *JAK2*, *IL6* y *RHBDD3* entre individuos con ECe y controles. La expresión de los dos últimos genes sólo está alterada en pacientes con debut en edad adulta, los cuales difieren de pacientes con debut en edad pediátrica. Esto parece indicar la existencia de un perfil de expresión diferencial según la edad de debut, que podría estar relacionada con las diferencias en la presentación clínica a menudo observadas entre pacientes celíacos en edad pediátrica o adulta. Estratificando nuestros pacientes con ECe por tratamiento observamos un aumento de expresión de *SMAD3* en los pacientes con dieta libre de gluten, aunque se necesita validar este resultado en nuevos estudios. Dado que la mayoría de los estudios de expresión en EII se centran en la EC, analizamos también la expresión de estos genes en pacientes con CU. Observamos mayor expresión de *CCR6*, *IL6* y *SMAD3* en muestras de tejido inflamado. Nuestros resultados en *CCR6* e *IL6* confirman los estudios previos mientras que en el caso del gen *SMAD3* son necesarios réplicas de nuestro estudio. El manuscrito que recoge estos datos está enviándose para su publicación.

5. Estudios de farmacogenética en pacientes con EC en tratamiento con IFX.

   Observamos una mayor frecuencia del haplotipo rs1061624_A-rs3397_T del gen *TNFRSF1B* en los individuos no respondedores al tratamiento con IFX (p=0,015, OR=1,78 95% IC 1,09-2,90), concordante con lo observado previamente en población japonesa. En el análisis de las características

fenotípicas de nuestros pacientes observamos una mejor respuesta en los pacientes con menos años de evolución, resultado ya previamente descrito en otros estudios y que hemos publicado recientemente en Hum Immunol. 2013 Oct 10. doi:pii: S0198-8859(13)00561-2. Ninguna de las características demográficas analizadas parece actuar como un factor de confusión.

Observamos una asociación con pérdida de respuesta a IFX de una combinación alélica en 4 genes estudiados (*TNFAIP6*, *IL11*, *G0S2*, *S100A8* y *S100A9*; p= 0,0098 OR=6,39 95% IC 1,32- 32,98) cuya expresión conjunta se vio alterada previamente en pacientes con EC. A diferencia del caso anterior, ninguna de las características demográficas muestra ningún tipo de asociación con la respuesta. Tampoco observamos ningún factor de confusión. Nuestros resultados parecen indicar que los cambios de expresión observados previamente están influenciados por la propia genética del individuo.

**Conclusiones**

Este trabajo define nuevos factores de susceptibilidad implicados tanto en EII como en ECe y puede ayudar a comprender los mecanismos moleculares involucrados en estas enfermedades. Además, define marcadores genéticos que predicen una pérdida de respuesta a IFX, uno de los tratamientos de elección en la EC. Las conclusiones de este trabajo son:

1. Existe un factor genético de susceptibilidad a ECe localizado en el HLA y presente en el haplotipo ancestral 8.1. El riesgo conferido por este factor sólo se observa cuando los individuos no portan un segundo alelo *DQB1*02*. Este factor de riesgo podría también estar presente en individuos DQ2.5 *trans*, dado el riesgo similar a padecer ECe que presentan ambos conjuntos de individuos.

2. La interacción entre variantes genéticas presentes en *TLR9* e *IL23R* contribuye al desarrollo de la enfermedad de Crohn en población blanca. Por el contrario, la interacción previamente descrita en población alemana entre *IL23R* y *NOD2* no parece influir.

3. Los principales genes implicados en la respuesta inmune tipo Th17 no parecen estar asociados al riesgo a padecer enfermedad celiaca en población española. Por el contrario, la expresión alterada de los genes *IL23A*, *JAK2*, *IL6* y *RHBDD3* en pacientes de ECe sugiere que la ruta Th17 está implicada en la patogenia de la enfermedad.

4. Dependiendo de la edad de debut de la enfermedad celiaca puede existir un perfil de expresión génica diferente, como sugieren nuestros resultados en *IL6* y *RHBDD3*.

5. La expresión de *CCR6*, *IL6* y *SMAD3*, todos ellos implicados en la respuesta tipo Th17, aumenta en regiones inflamadas de pacientes con colitis ulcerosa activa.

6. La ruta Th17 parece importante en la patogenia de la enfermedad celiaca y colitis ulcerosa como se refleja por la expresión alterada de ciertos genes que, sin embargo, parece diferir entre ambas enfermedades.

7. El gen *TNFRSF1B* está implicado en la respuesta a infliximab en individuos con enfermedad de Crohn.

8. La expresión diferencial de los genes *TNFAIP6*, *S100A8*, *IL11*, *G0S2* y *S100A* observada entre pacientes de enfermedad de Crohn respondedores y no respondedores a infliximab parece venir determinada por variantes específicas presentes en esos genes.

**Introduction**

Inflammatory bowel disease (IBD) and celiac disease (CeD) are two immune-mediated disorders with chronic intestinal affectation. In CeD, it is well known the etiological role of gluten in disease development, while in IBD commensal bacteria could be one of the causes involved in the altered immune response.

In the last years large numbers of variants have been associated with these diseases by Genome Wide Association Studies (GWAS): 163 loci with IBD and 41 with CeD. Despite the high number of new variants, it exists a percentage of heritability remains to be discovered ("missing heritability") which seems to be a consequence of different factors as gene-gene or environment-genes interactions. Moreover, specific genetic factors can affect specific groups of patients, which has not been studied in GWAS.

In CeD, the treatment consists in a gluten free diet (GFD), which is effective in almost all patients. On the contrast, in IBD a wide variety of treatments with different efficacy exist. Chron´s disease (CD) patients are administrated infliximab (IFX) (monoclonal antibody anti-TNF-α) when conventional treatments are not efficient, but the response to IFX depend on different factors, which probably include genetic factors.

**Hypothesis**

- Gene-Gene interactions can explain at least part of the unexplained genetic variance, the so named "missing heritability", in complex diseases as CeD and IBD.
- A dysbalance of the Th17 pathway can contribute to CeD development. This idea is not supported by GWAS results but gene-gene interactions have not been included in GWAS.
- Intestinal expression studies of genes associated with CeD and ulcerative colitis (UC) susceptibility can help to understand how these diseases develop.
- The individual genetic background can affect the response to IFX.

**Aims**

- ✓ To search for genetic factors associated with celiac disease and inflammatory bowel disease which explain the "missing heritability".
  - To search for new CeD susceptibility variants in the HLA region.
  - To confirm the gene-gene interactions (*IL23R*, *CARD15* and *TLR9*) as risk factors in inflammatory bowel disease.
  - To study the association of genes involved in the Th17 pathway with the risk to develop CeD.
- ✓ To study the intestinal expression of genes involved in Th17 pathway in CeD and UC.
- ✓ To study risk factors which predict the response to infliximab in CD patients.

**Material and Methods**

1. Study of new susceptibility variants in the HLA region in CeD.

   We studied 274 trios (the affected individual and both parents), 369 CeD patients and 461 individuals without immune mediated diseases as controls. *HLA-DRB1*, *-DQA1*, *-DQB1, -B8* genes, *TNFa* and *TNFb* microsatellites and *TNF* -308 (rs1800629) and -376 (rs1800750) polymorphisms were studied to know the presence of the two main haplotypes which contain the *DRB1*03:01* allele in the studied individuals. Familial data were analyzed with the transmission disequilibrium test (TDT) and case-control data with chi-square tests or the Fisher´s exact test when some expected value was below 5.

2. Gene-gene interactions (*IL23R*, *CARD15* and *TLR9*) as risk factors in inflammatory bowel disease.

   We studied a first sample of 416 CD patients, 452 UC patients and 547 healthy individuals and a second sample of 122 CD patients and 307 controls. Nine single nucleotide polymorphisms (SNPs) in *CARD15*, *IL23R* and *TLR9* were analysed by TaqMan technology. Association studies were performed by chi-square tests or the Fisher´s exact test.

3. Association study of genes involved in the Th17 pathway with CeD susceptibility.

   A first sample of 735 CeD patients and 549 healthy individuals and a second sample of 294 CeD individuals and 475 controls were studied. A total of 101 SNPs located in 16 genes were analyzed (*CCR6*, *GLB1*, *IL17F*, *IL2*, *IL21*, *IL22*, *IL23A*, *IL23R*, *IL6*, *JAK2*, *RHBDD3*, *SMAD3*, *STAT3* and *TYK2*) by Veracode technology. The SNPs were selected to cover most of the variability of the studied regions by aggressive tagging in all genes except *IL6R, JAK2* and *STAT3*. The case-control association study was performed by chi-square tests or the Fisher´s test and epistatic interactions were evaluated by logistic regression, random forests, classification and regression trees and multifactor dimensionality reduction.

4. Intestinal expression study of genes involved in the Th17 pathway in CeD and UC.

   We studied the expression levels of 12 Th17 related genes (*CCR6*, *IL17F*, *IL2*, *IL21*, *IL22*, *IL23A*, *IL23R*, *IL6*, *JAK2*, *RHBDD3*, *SMAD3* and *STAT3)* in 24 duodenal biopsies from CeD patients and 6 from healthy individuals by quantitative PCR. We stratified the biopsies depending on age at onset (14 pediatric and 6 adult patients) and treatment (4 patients on a gluten free diet and 20 on gluten containing diet). The geNorm version 3.5 and NormFinder version 0.953 programs were used to determine the reference genes and *GLB1* and *TYK2* genes were selected. Expression levels were determined by the ΔΔCt method ($\Delta Ct_{\text{target gene}} - \Delta Ct_{\text{reference gene}}$). Expression level comparisons between groups were performed by the U Mann Whitney test using SPSS 15.1 .

   The expression levels of those genes were also analyzed in biopsies from colon and/or rectum from 8 UC patients. A sample from healthy and inflamed tissue was analyzed. Statistic analyses were performed by the Wilcoxon test (paired samples) using the SPSS software.

5. Study of risk factors to predict the response to infliximab of Chron´s disease patients.

A first pharmacogenetic study included 238 CD IFX responders and 59 non-responders. Three doses of IFX were administrated to patients at week 0, 2 and 6 and the response to treatment was evaluated at week 10 after the first dose by the Harvey-Bradshaw index (HBI). Partial response was considered as HBI<3 and remission as HBI≤4. Four SNPs in *TNFRSF1B* (rs1061622, rs1061624 y rs3397) and *TNFRSF1A* (rs767455) were studied by TaqMan technology. Haplotypes between SNPs in linkage disequilibrium in *TNFRSF1B* were obtained with the Haploview software 4.1 using the Expectation-Maximization algorithm. Allele and haplotypic frequencies were compared between responders and nonresponders by chi-square tests and the Fisher´s test. Demographic characteristics patients as smoking, Montreal classification, years of disease, concomitant treatment… were evaluated as confounders variants using logistic regression with the SPSS software.

A second pharmacogenetic study was carried out studying five genes (*TNFAIP6*, *S100A8*, *IL11*, *G0S2* and *S100A9*) in 210 IFX treated CD patients (166 responders and 44 non-responders). Administrated doses, response to IFX and statistic analyse were designed as in the previous study.

**Results and discussion**

1. Study of new susceptibility variants in the HLA region in CeD.

   We observed an additional risk variant in the ancestral haplotype (AH) 8.1, which is present in some DQ2 patients. This variant seems to confer additional risk only to individuals who do not carry a second *DQB1*02* allele. This factor could be also present in DQ2.5 *trans* individuals because both groups present a similar risk to develop CeD. Using data from the Immunochip project we searched for a shared variant between AH 8.1 and DQ2.5 *trans* individuals, but we did not find any variant, althought the statistical power of our study is compromised due to the high complexity of the HLA region. Our results have been published: *HLA and celiac disease susceptibility: new genetic factors bring open questions about the HLA influence and gene-dosage effects.* PLoS One. 2012;7(10):e48403

2. Gene-gene interactions (*IL23R*, *CARD15* and *TLR9*) as risk factors in IBD.

Our results show that the minor allele of rs7517847 in *IL23R* diminishes CeD risk in individuals *TLR9* rs352162_CC (p=0.0003 OR=0.44 95% CI 0.27-0.71). Therefore, we observed an epistatic interaction between *IL23R* and *TLR9* genes in Crohn's disease (CD) which was previously found in German population. On the contrary, we did not observe interaction between the polymorphisms in the *IL23R* and the *CARD15* genes. These results are interesting because of the functional implication of the coding protein in pathogen recognition and defense. A paper with these results has been submitted for publication.

3. Association study of genes involved in the Th17 pathway with CeD susceptibility.

   We did not observe any significant association when studing the 101 selected SNPs or the epistatic interactions between the studied genes (*CCR6*, *GLB1*, *IL17F*, *IL2*, *IL21*, *IL22*, *IL23A*, *IL23R*, *IL6*, *JAK2*, *RHBDD3*, *SMAD3*, *STAT3* and *TYK2*), to know the implication of the Th17 pathway. We confirm that no exist genetic implication of the Th17 pathway in CeD development exist. We show these results in the published article: *Th17-related genes and celiac disease susceptibility.* PLoS One. 2012;7(2):e31244.

4. Study of the intestinal expression of genes involved in the Th17 pathway in CeD and UC

   We observed a different expression profile in four genes (*IL23A*, *JAK2*, *IL6* and *RHBDD3*) between CeD and controls. The different expression observed for *IL6* and *RHBDD3* was only present in CeD patients with adult onset. The commonly observed differences in the clinical presentation between children and adults could be partially due to differences in the inflammatory response. Evaluation of the expression of these Th17 related genes in CeD patients on a GFD showed an increase of *STAT3* intestinal expression in these patients compared to those on gluten containing diet, but it is necessary to validate our results. In IBD, most of the intestinal gene expression studies have been performed in CD patients, therefore we analyzed the expression of Th17 genes in the less studied UC patients. Increased levels of *IL6*, *CCR6* and *SMAD3* in inflamed tissue were observed. Our data confirm that *IL6* and *CCR6* play a role in mediating

inflammation in CU. Further studies are necessary to validate our results in *SMAD3*. A paper with these results has been submitted for publication.

5. Study of risk factors to predict the response of infliximab in CD patients.

   We observed that the haplotype rs1061624_A-rs3397_T in the *TNFRSF1B* gene was significantly increased in nonresponders to IFX (p=0.015, OR=1.78, 95% CI 1.09-2.90). Our work confirms the association of the *TNFRSF1B* gene in the loss of response to IFX, which was previously found in Japanese population. The comparison of clinical characteristics between responders and nonresponders to IFX showed a significant difference for disease duration, as previously described in the literature. Demographic characteristics were tested as potential confounders but none of them showed any effect. The results have been published in Hum Immunol. 2013 Oct 10. doi:pii: S0198-8859(13)00561-2.

   In addition, we observed one allelic combination of the *TNFAIP6*, *S100A8*, *IL11*, *G0S2* y *S100A9* genes associated with loss of response to IFX. The expression levels of these genes predict responsiveness in colonic CD patients in a previous study in white population. Our results seem to indicate that those differences of gene expression are influenced by an individual genetic predisposition.

**Conclusions**

This work associates new risk factors in IBD and CeD and it can help to understand the molecular mechanisms involved in the disease. In addition, we found genetic markers which predict loss of response to IFX, one of the elections to treat CD patients. The conclusions of this work are:

1. There is a risk variant for celiac disease in the HLA region which is present in the ancestral haplotype 8.1. The risk conferred by this factor is only observed in individuals who do not carry a second allele *DQB1*02*. This factor can be also present in DQ2.5 *trans* individuals, due to the similar CeD risk in both groups.

2. The interaction between genetic variants in *TLR9* and *IL23R* plays a role in Crohn´s disease susceptibility in white population. By the contrary, the previously described interaction in German population between *IL23R* and *NOD2* does not seem to influence disease risk.

3. The main genes involved in the Th17 pathway do not seem to influence celiac disease risk in Spanish population. By the contrary, our results show the altered expression of *IL23A*, *JAK2*, *IL6* and *RHBDD3* in celiac disease patients, suggesting that the Th17 pathway is involved in CeD pathogeny.

4. Depending on the age at onset of celiac disease a different expression profile may exist, as our results in *IL6* and *RHBDD3* genes suggest.

5. The expression levels of the *CCR6*, *IL6* and *SMAD3* genes, which are involved in the Th17 immune response, are elevated in inflamed regions or ulcerative colitis patients.

6. The Th17 pathway can be important in the pathogeny of celiac disease and ulcerative colitis as a result of the altered expression observed in several genes, which seem to differ between both diseases.

7. The *TNFRSF1B* gene is involved in the response to infliximab of Crohn's disease patients.

8. The observed differential expression of the *TNFAIP6*, *S100A8*, *IL11*, *G0S2* y *S100A* genes between Crohn´s disease responders and non-responders to infliximab seems to be a consequence of specific genetic variants in those genes.

## 1. Aspectos generales

La enfermedad inflamatoria intestinal (EII) y la enfermedad celiaca (ECe) se caracterizan por ser enfermedades inflamatorias crónicas intestinales. Aunque se desconoce su etiología, en ambos casos se trata de enfermedades complejas caracterizadas por la implicación tanto de factores genéticos como ambientales.

### 1.1 Enfermedad inflamatoria intestinal

La enfermedad inflamatoria intestinal (EII) se caracteriza por una respuesta inflamatoria inapropiada frente a la flora intestinal en individuos genéticamente predispuestos. La importancia de las interacciones bacteria-hospedador en la patogenia de la enfermedad se ha puesto en evidencia en diversos estudios. [1, 2]

La EII presenta principalmente dos formas clínicas: enfermedad de Crohn (EC) y colitis ulcerosa (CU). En la EC está afectada principalmente la zona del íleon y colon pero pueden aparecer daños en cualquier zona del tracto digestivo, mientras que en la CU el daño está más localizado en la parte final del tracto digestivo. En el caso de la EC existe una afectación continua y transmural con posible aparición de fístulas, mientras que en el caso de la CU se aprecia un patrón discontinuo con afectación de la mucosa y sin aparición de fístulas o granulomas.

La prevalencia de la EC y CU está entre 26-99 casos por cada 100.000 habitantes y entre 37-246 casos por cada 100.000 habitantes, respectivamente, [3, 4] siendo mayor en países desarrollados y áreas urbanas. La incidencia es mayor en ambos casos entre la primera y tercera década de vida [5] y no parece que difiera entre hombres y mujeres. [6]

Es frecuente la aparición de síntomas extraintestinales en pacientes con EII como artralgias, aftas orales, eritema nodoso, osteopenia, uveitis…,[7] así como la presencia concomitante de otras enfermedades inflamatorias o autoinmunes como colangitis esclerosante, espondilitis anquilosante y psoriasis [3].

El comportamiento de la enfermedad y la localización del daño es variable en la EC, lo que permite una buena clasificación de los pacientes. En la actualidad se utiliza la clasificación de Montreal, [8] la cual se basa en la edad de diagnóstico, localización, comportamiento de la enfermedad y posible afectación perianal del paciente ( tabla 1).

Tabla 1. Clasificación de Montreal en enfermedad de Crohn.

| | |
|---|---|
| Edad de diagnóstico (A) | |
| A1 | < 16 años |
| A2 | 17-40 años |
| A3 | > 40 años |
| Localización (L) | |
| L1 | Íleon terminal |
| L2 | Colon |
| L3 | Íleo-colónico |
| L4 | Tracto superior |
| L1+L4 | Íleon terminal+ Tracto superior |
| L2+L4 | Colon+ Tracto superior |
| L3+L4 | Íleo-colónico+ Tracto superior |
| Comportamiento (B) | |
| B1 | No estenosante, no penetrante |
| B2 | Estenosante |
| B3 | Penetrante |
| B1+p | No estenosante, no penetrante + perianal |
| B2+p | Estenosante + perianal |
| B3+p | Penetrante + perianal |

En el caso de la CU también se emplea la clasificación de Montreal, la cual se basa en la extensión y gravedad de la enfermedad (tabla 2). [8]

Tabla 2. Clasificación de Montreal en colitis ulcerosa.

| | |
|---|---|
| Extensión (E) | |
| E1 | Proctitis Ulcerosa (limitado a la zona del recto) |
| E2 | Colitis izquierda o distal (limitado al colon izquierdo sin superar el ángulo esplénico) |
| E3 | Colitis extensa o Pancolitis (afectación que se extiende mas allá del ángulo esplénico) |
| Gravedad (S) | |
| S0 | Colitis en remisión |
| S1 | Colitis leve |
| S2 | Colitis moderada |
| S3 | Colitis grave |

### 1.2 Enfermedad celiaca

La enfermedad celiaca (ECe) se caracteriza por ser una enfermedad crónica intestinal que aparece en individuos genéticamente predispuestos tras la ingesta de gluten, denominado así el conjunto de proteínas presentes en el trigo (gliadinas y gluteninas) y sus homólogos en la cebada y centeno, con un alto contenido en glutamina y prolina.[9]

La prevalencia de la ECe en población blanca es aproximadamente del 1%,[10, 11] con un pico de incidencia en la primera década de vida. La incidencia en la enfermedad es mayor en mujeres que en hombres (2:1), auque esta diferencia parece desaparecer con la edad [12]. Estudios recientes en población española muestran una prevalencia menor, de 1:200, viendo diferencias entre adultos (1:357) y niños (1:71) [13]. La prevalencia de la ECe es mayor en ciertos grupos de riesgo, donde se incluyen familiares de primer grado de individuos con ECe e individuos con otro tipo de patologías como la DM1 (diabetes de tipo 1),[14] tiroiditis autoinmune [15] y deficiencia de IgA.[16]

Sólo el 10-15% de la población celiaca es diagnosticada y tratada. [17] A pesar de los esfuerzos en mejorar el diagnóstico de la enfermedad, fundamentalmente basados en la búsqueda de mejores marcadores serológicos, sigue habiendo un porcentaje alto de personas no diagnosticadas, debido principalmente a la ausencia de sintomatología. La epidemiología de la ECe se compara con frecuencia con un iceberg, los pacientes con sintomatología constituyen la mayoría de los casos sobre la superficie del mar.[18]

La presentación clínica de la ECe es muy variada. Según las últimas publicaciones [19] se define ECe clásica como la caracterizada por signos o síntomas de malabsorción, como pérdida de peso, retraso del crecimiento, diarrea y esteatorrea. Mientras que en el caso de ECe "no clásica", la sintomatología no incluye signos de malabsorción y la clínica puede ser muy variada (aftas orales, infertilidad, problemas neurológicos, irritabilidad, fatiga crónica…). Existe un tercer grupo de pacientes que son los clasificados como celíacos potenciales, presentan una mucosa normal del intestino delgado y un aumento del riesgo a padecer la enfermedad debido a una serología específica positiva. Existen diferencias de presentación entre adultos y niños, predominando la forma clásica en niños y la forma "no clásica" en los adultos.

## 2. Factores ambientales

### 2.1 Factores ambientales en EII

La patogénesis de la EII se ha relacionado con la interacción de la mucosa intestinal y bacterias patógenas y/o comensales de la flora intestinal.[20, 21] La hipótesis de la higiene se ha postulado como una posible causa de enfermedades inflamatorias autoinmunes, entre ellas EII, la enfermedad se produciría como resultado de una pérdida de exposición temprana a agentes microbianos debido a un exceso de limpieza en el medio. [22]

Bacterias como *Escherichia coli*, *Listeria monocytogenes*, *Yersinia enterocolitica* y *Mycobacterium avium paratuberculosis* se han visto aumentadas en la mucosa intestinal de pacientes de EII.[23, 24] Otros microorganismos de la especie *Helicobacter* se han asociado con la CU. Otros estudios también observaron una alteración de la flora bacteriana en el medio intestinal, en concreto una disminución de bacterias anti-inflamatorias como es *Faecalibacterium prausnitzii* en el íleon de

pacientes con EC, y un aumento de los niveles de *Escherichia coli* en la mucosa ileal de pacientes con EII, siendo esta última bacteria capaz de inducir la producción de TNF-α, una citocina inflamatoria. [25, 26]

Otros factores ambientales como fumar y la apendicectomía parecen influir en el riesgo de padecer EII. En el caso del tabaco tiene un efecto opuesto en las dos enfermedades: beneficioso en CU y dañino en EC.[27, 28] Una apendicectomía antes de los 20 años reduce el riesgo de desarrollar CU, una posible extirpación del apéndice puede alterar el balance entre células T efectoras y T reguladoras, ya que en el apéndice existe una agregación de nódulos linfáticos. [29]

Los niveles de vitamina D, la dieta, el uso de hormonas anticonceptivas y el estrés son algunos de los factores que también se han asociado con el riesgo a padecer EII, pero son necesarios más estudios para confirmar su papel.[30]

### 2.2 Factores ambientales en ECe

El principal factor ambiental en la ECe es la ingesta de gluten en la dieta, necesario para el desarrollo de la enfermedad. Además puede haber otros factores que contribuyan al riesgo, así también se han relacionado infecciones como desencadenantes de la enfermedad, como infecciones virales por *Rotavirus*. [31] Las infecciones virales pueden producir niveles elevados de IL-15, estimulando un proceso inflamatorio. En el caso de pacientes refractarios, los cuales no responden al tratamiento de una dieta libre de gluten, se ha observado un aumento de infecciones con ciertos virus como son el virus de la hepatitis B y C. [32]

## 3. Factores genéticos

### 3.1 Abordaje de los estudios genéticos

Los factores genéticos juegan un papel importante en el desarrollo de las enfermedades autoinmunes, pero su estudio entraña una gran dificultad debido a la existencia de un componente genético heterogéneo y a la baja penetrancia de la mayoría de los alelos de riesgo, de modo que cada alelo implicado contribuye con un efecto muy pequeño.

Existen diferentes modos de abordar el estudio del componente genético; en todos ellos cuanto mayor sea el número de individuos estudiados existe mayor potencia para obtener resultados significativos. Puesto que se pueden obtener resultados falsos positivos, es de gran importancia confirmar un resultado ya sea en una muestra independiente de la misma población o en distintas poblaciones para además conocer su implicación generalizada en la enfermedad.

Dos son los principales tipos de estudio utilizados en el abordaje de la búsqueda de marcadores genéticos:

Estudios de ligamiento llevados a cabo en familias con diversos miembros afectados. Han empleado microsatélites o polimorfismos de un solo nucleótido (SNPs) como marcadores para determinar si los individuos afectados comparten información genética en alguna región más frecuentemente de lo esperado.[33] De esta manera se han identificado regiones amplias de ligamiento en EII (hasta 9 loci) [34, 35] y en ECe (hasta 10 loci) [36, 37], con varios genes presentes en cada región. A pesar del amplio esfuerzo, estos estudios apenas han identificado variantes de riesgo, las principales razones del bajo número de variantes identificadas con este método son la baja penetrancia de las variantes causales, haciendo más difícil detectarlas por cosegregación en un pedigree, y la hetereogenidad entre las familias. [38]

Estudios de asociación caso-control, se basan en la comparación entre un grupo caso (enfermos) frente a un grupo control (sanos) de una misma población. Buscan factores genéticos cuya frecuencia varíe entre ambos grupos. En las últimas décadas se han empleado principalmente SNPs, seleccionados en base a una información previa como puede ser su función conocida en la enfermedad, estudios previos en modelos múridos o estar en regiones ligadas en estudios familiares.

En los últimos años se han llevado a cabo estudios caso-control a gran escala. Estos estudios de barrido genómico (genome wide association studies, GWAS) carecen de hipótesis previa y emplean cientos de miles de marcadores genéticos distribuidos por todo el genoma para tratar de recoger gran variabilidad en todas las regiones. Los primeros estudios fueron llevados a cabo por el *Wellcome Trust Case-Control Consortium* (WTCCC), gracias al empleo de la base de datos originada en el proyecto HapMap e identificaron decenas de SNPs a lo largo de todo el genoma asociados a diversas enfermedades autoinmunes (DM1, AR o EC).[39] Posteriormente

se han llevado a cabo numerosos estudios, en los que ha ido aumentando el número de variantes incluidas en parte debido a la descripción de nuevas variantes genéticas gracias al proyecto de los 1000 genomas y al mejor conocimiento de la estructura genética de las poblaciones. Los estudios de GWAS aprovechan la existencia de desequilibrio de ligamiento (LD), el cual implica que alelos de dos loci se heredan a la vez más frecuentemente de lo que se esperaría por azar. A través del LD se puede obtener información de variantes genéticas no directamente analizadas (imputación). Como resultado, estos estudios contemplan la asociación de un número elevadísimo de marcadores genéticos. Se habla de marcadores puesto que en la gran mayoría de los casos se obtienen señales asociadas a la enfermedad, que son señales respecto a una variante causal en LD. Estos estudios han identificado multitud de SNPs asociados a diversas enfermedades complejas, entre ellas numerosas enfermedades mediadas por el sistema inmunológico. Además han mostrado diversos marcadores asociados a distintas enfermedades, sugiriendo un mecanismo molecular común.[40]. Sin embargo, el componente genético de cada enfermedad está lejos de conocerse en su totalidad dado que este tipo de estudios presenta ciertas limitaciones.

### 3.2 Genética en la EII

La concordancia entre gemelos monocigóticos es de 10-15% en CU frente al 30-35% en EC, sugiriendo que los factores genéticos tienen menor importancia en CU que en EC.[41]

Inicialmente los estudios de ligamiento permitieron identificar hasta nueve regiones ligadas a la EII. Tres regiones denominadas *IBD1*, *IBD3* e *IBD5* se han replicado consistentemente en la enfermedad. [34, 35]

La región *IBD1* se localiza en el cromosoma 16 e incluye el gen *NOD2*, uno de los genes que confiere mayor riesgo a padecer EC y el primero asociado a esta enfermedad.[42, 43] Tres son las mutaciones descritas en este gen (Arg702Trp, Gly908Arg y Leu1007fsinsC). Un 30% de los pacientes con EC presenta al menos una de estas mutaciones, [44, 45] respecto a un 10% de la población general. En población europea se observó un efecto de dosis en este gen: individuos que portan sólo una de estas mutaciones presentan de 2 a 4 veces más riesgo a desarrollar EC, e individuos con dos mutaciones o una en homocigosis presentan de 20 a 40 veces más riesgo. [46-48]

La región *IBD3* se encuentran localizada en el brazo corto del cromosoma 6 (6p21), donde se encuentran los genes HLA (*Human Leukocyte Antigen*). Se trata de una región de genes altamente polimórficos y donde existe elevado LD. Existe una mayor implicación del HLA en CU que EC, aunque esta región juega un papel mucho menos relevante que en otras enfermedades autoinmunes como ECe o DM1. En CU se asocia el alelo *HLA DRB1*01:03*, [49, 50] también asociado en menor medida al riesgo a desarrollar EC con afectación colónica. [51-53] El locus *IBD5* se localiza en la región 5q31 e incluye un grupo de genes que codifican citocinas (IL-4, IL-5, IL-13), factores de regulación (*IRF1)* y genes que codifican transportadores catiónicos (*OCTN1/SLC22A4*, *OCTN2/SLC22A5*).

Posteriormente, los estudios de GWAS en la EII han tenido un gran éxito en el número de variantes asociadas, con 163 loci identificados, 110 compartidos por ambas enfermedades, y 30 y 23 loci asociados independientemente a EC y CU, respectivamente [54-56]. A pesar del elevado número de variantes descritas, la varianza genética explicada hasta el momento es del 13,6 % en EC y 7,5% en CU.[56] A pesar de ser enfermedades con características clínicas e inmunopatológicas diferentes, los estudios genéticos señalan mecanismos etiopatológicos comunes.

Recientemente, un análisis de GWAS en población española ha determinado un nuevo locus asociado a EC en la región 22q13.2, demostrando que el estudio en poblaciones distintas puede ser útil para identificar nuevos loci de susceptibilidad.[57]

### 3.3 Genética en la ECe

En la ECe, los factores genéticos contribuyen de manera importante al desarrollo de la enfermedad como lo muestra la concordancia entre gemelos monocigóticos que es del 75%, mientras que en gemelos dizigóticos es del 11%. [58]

Inicialmente se llevaron a cabo estudios de ligamiento en los que se identificaron hasta 10 regiones diferentes ligadas a la ECe. [36, 37] Sin embargo, únicamente cuatro fueron identificadas en diversos estudios de ligamiento: *CELIAC1*, la cual incluye la región del HLA; *CELIAC2*, localizada en 5q31 y también ligada a EII (*IBD5*); *CELIAC3*, localizada en 2q33 la cual incluye genes reguladores (*CD28*, *CTLA4* e *ICOS*); y *CELIAC4*, en 19p13.

En la ECe, el factor principal de riesgo genético a padecer la enfermedad reside en el HLA. Además de observar una fuerte asociación en esta región, los primeros

estudios de GWAS determinaron hasta 26 loci no-HLA asociados a la enfermedad [59-62] y estudios posteriores incluyendo mapeo fino un total de 39 loci asociados a la ECe. [62] Un estudio reciente de secuenciación del exoma de 25 regiones asociadas previamente en GWAS añadió un nuevo gen asociado a la ECe, el gen *NCG2*, que codifica un componente del complejo respiratorio NADPH oxidasa de los neutrófilos.[63] El cómputo global de los estudios llega a explicar un 54% de la variabilidad genética incluyendo tanto loci en el HLA (40%) como los loci no-HLA (13.7%).[62]

Prácticamente el 100% de los pacientes de ECe presentan la molécula HLA DQ2 o HLA DQ8. [64] La molécula DQ2 es un heterodímero formado por una cadena β codificada por el alelo *HLA DQB1*02* (*DQB1*02:02* o *DQB1*02:01*) y una cadena α codificada por el alelo *HLA DQA1*05* (*DQA1*05:01* o *DQA1*05:05*), denominando al conjunto DQ2.5. En torno a un 90-95% del total de pacientes presentan la molécula DQ2, mientras que del 5 al 10% restante presentan la molécula DQ8, también un heterodímero codificado por una cadena β y una α codificadas por los alelos *HLA-DQB1*03:02* y *HLA-DQA1*03*, respectivamente. La molécula DQ2 puede heredarse en configuración *cis* (cuando procede de un progenitor) o *trans* (cuando procede de ambos progenitores) (Figura 1). En el caso de la configuración *cis*, a su vez presenta el alelo *DRB1*03*, debido al elevado desequilibrio de ligamiento que presenta la región.

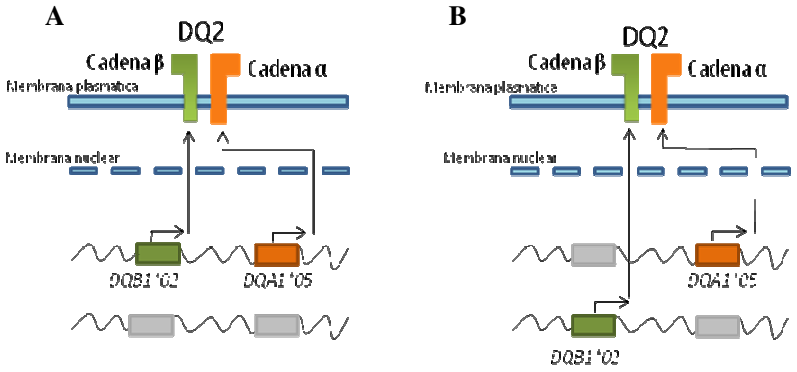



Fig 1. Representación de la molécula DQ2 heredada en *cis* (A) y en *trans* (B).

Las moléculas de HLA confieren riesgo a ECe con un efecto de dosis, dependiendo del número de alelos *DQB1*02* y *DQA1*05* presentes, el individuo tendrá un riesgo alto o intermedio de padecer la enfermedad (Tabla 3).[65, 66] Individuos homozigotos DQ2.5*cis* o heterozigotos DQ2.5*cis* con un segundo alelo *DQB1*02* (DQ2.2) presentan mayor riesgo a padecer la ECe. Mientras que en aquellos individuos con una sola copia de la molécula DQ2 tanto si está codificada de manera *cis* o *trans* presentan un riesgo intermedio. Aquellos individuos con la molécula DQ8 presentan un riesgo intermedio-bajo y al igual que con la molécula DQ2, existe un efecto de dosis en el riesgo conferido por DQ8. [67] Por último aquellos individuos que presentan una copia del alelo *HLA-DQA1*02* (DQ2.2) tienen riesgo bajo. [67]

**Tabla 3:** Genotipos HLA asociados a celiaquía.

| **Genotipo DQ** | **Tipo DQ** | **Riesgo** |
|---|---|---|
| *DQB1*02 DQA1*05/ DQB1*02 DQA1*05* | DQ2.5 homocigotos | Muy alto |
| *DQB1*02 DQA1*05/ DQB1*02 DQA1*02:01* | DQ2.5 *cis*+ DQ2.2 | Muy alto |
| *DQB1*02 DQA1*05/-* | DQ2.5 *cis* heterocigoto+ no DQ2.2 | Alto |
| *DQB1*02 -/- DQA1*05* | DQ2.5 *trans* | Alto |
| *DQB1*03 DQA1*03:01/ DQB1*03 DQA1*03:01* | DQ8 homocigotos | Alto |
| *DQB1*03 DQA1*03:01/-* | DQ8 heterocigoto | Moderado |
| *DQB1*02 DQA1*02:01/-* | DQ2.2+ no-DQ2.2/no DQ8 | Bajo |
| -/- | Resto | Sin riesgo |

- representa un alelo distinto de *DQA1*05* o *DQB1*02*.

La diferencia de riesgo en los diferentes grupos reside en la capacidad de unión con los péptidos derivados de gluten por parte de la molécula DQ. La molécula HLADQ2.5 puede presentar un mayor repertorio de péptidos derivados de gluten que

la molécula HLADQ2.2, además existe mayor avidez en la unión. [68] En la presentación de moléculas DQ8 también parece que ciertos péptidos de gliadina parecen dominar, pero son péptidos más fácilmente degradables y que requieren más pasos de deamidación que en el caso de la molécula DQ2.5, con lo que limita más la presentación del antígeno y por esa causa disminuye el riesgo a desarrollar la ECe en los portadores de DQ8. [69]Por lo tanto, la exposición de gluten depende tanto del tipo como de la cantidad de moléculas HLA-DQ, que es lo que determina la eficiencia en la presentación de los péptidos de gluten a las células T CD4+. [70]

De ahí que el genotipado de los alelos que codifican la molécula HLA-DQ represente uno de los primeros pasos para el estudio de familiares de primer grado de pacientes con ECe y de individuos que pertenecen a grupos de riesgo.[71] El no presentar ninguna de las combinaciones alélicas del HLA asociadas se considera excluyente para desarrollar la enfermedad, aunque existe un porcentaje minoritario de casos que no presentan el HLA asociado a la ECe y desarrollan la enfermedad. [67]

La combinación alélica *DRB1*03*, *DQA1*05* y *DQB1*02* suele aparecer con otros alelos que se han mantenido altamente conservados en el tiempo que es lo que se denomina Haplotipo Ancestral (AH). Dentro del contexto de *DRB1*03* hablamos del AH 8.1 que incluye los alelos *DRB1*03:01, DQA1*05:01, DQB1*02:01, TNF -308A, TNFa2b3* y *B*8*; y del AH18.2 que incluye los alelos *DRB1*03:01, DQA1*05:01, DQB1*02:01, TNF -376A* , *TNFa1b5* y *B*18*. Los haplotipos se denominan dependiendo del alelo en *HLA-B* que presenten.

En ciertas enfermedades de base inmunológica como son DM1 o EM se ha visto que el riesgo a padecer la enfermedad varía dependiendo del AH que presente el individuo. [72, 73]

Aproximadamente un 30% de la población blanca de ascendencia europea presenta la molécula DQ2, pero la prevalencia de la enfermedad en la población es del 1%. Por lo tanto, no todas las personas que presenten la molécula DQ2 o DQ8 e ingieran productos con gluten desarrollan la enfermedad. Otros factores genéticos y ambientales tienen que estar involucrados en la susceptibilidad a padecer ECe.

### 3.4 Búsqueda de la heredabilidad perdida

La heredabilidad indica la proporción de la varianza fenotípica que se debe a causas genéticas. A pesar de los grandes avances en la genética de las enfermedades

complejas, los estudios de GWAS explican un porcentaje pequeño de heredabilidad, entre un 20-30% en la mayoría de las enfermedades y en pocos casos más del 50%. Hasta ahora los GWAS explican un 23% y 16% de heredabilidad en EC y CU respectivamente, [54, 55] y en ECe incluyendo el HLA un 54%. [74]

Los GWAS presentan limitaciones que hacen necesario desarrollar nuevas alternativas de estudio. Hasta ahora los marcadores asociados presentaban una frecuencia mayor del 5%, una baja penetrancia y un efecto moderado (OR<1,5). Para alcanzar potencia estadística suficiente y detectar marcadores con una baja frecuencia se necesita un número muy elevado de pacientes y controles, por ello se reúnen individuos de diferentes poblaciones, pero esto dificulta la detección de variantes específicas de una población. A esto hay que añadir que el LD en algunas de las regiones asociadas no hace posible determinar cuál es la verdadera variante causal y en ocasiones tampoco cuál es el gen implicado, especialmente cuando los marcadores asociados se localizan en regiones con un alto contenido genético. Por lo tanto, son necesarios estudios genéticos y funcionales adicionales en esas regiones.

Estudios recientes se han basado en la secuenciación del genoma y el exoma pero apenas han determinado variantes nuevas que expliquen parte de esa "heredabilidad perdida", [63, 75] por lo tanto cabe pensar que existe alto porcentaje de "heredabilidad perdida" que ha de ser abordada con otro tipo de estudios.

Uno de los factores que podría estar involucrado en el porcentaje de heredabilidad sin conocer son las interacciones genéticas o epistasis. El efecto epistático se define como el efecto de enmascaramiento que posee una variante (locus o alelo) de manera que no se manifieste el fenotipo de una segunda variante.[76] Algunos autores apuntan que la "heredabilidad perdida" no se debe a una falta de variables por descubrir, sino a una sobreestimación de la heredabilidad total, consecuencia en parte de interacciones gen-gen. De ahí que Zuk et al. [77] indiquen que se deben llevar a cabo estudios de epistasis para ver hasta qué nivel influyen en la heredabilidad de la enfermedad. Puede que en ocasiones la interacción no sea exclusivamente entre dos genes sino entre un conjunto de genes, lo que hace más difícil su estudio.[78] Un ejemplo de interacción epistática en los estudios de GWAS es la observada para el gen *ERAP1* (aminopeptidasa 1 del retículo endoplásmico) involucrado en la presentación de los péptidos de HLA clase I. Tanto en espondilitis anquilosante como en psoriasis sólo se observa la asociación del *ERAP1* en los pacientes que presenten simultáneamente el alelo de susceptibilidad *HLA-B27* y *HLA-*

*Cw6*, respectivamente, con lo que a su vez sugiere un posible mecanismo común que involucra en las moléculas del *HLA-B27* y *HLA-Cw6* en ambas enfermedades.[79, 80]

Otros factores interesantes en la búsqueda de heredabilidad son factores ambientales como fumar, estrés o alimentación [81] y modificaciones epigenéticas en las que se engloban metilación, acetilación y fosforilación. [81]

Probablemente parte de la heredabilidad perdida resida en interacciones gen-gen y gen-ambiente. Sin embargo, su estudio entraña gran dificultad debido a la multitud de factores que hay que tener en cuenta y que pueden variar de unas poblaciones a otras.

### 3.5 Estudios de expresión génica

En muchos casos los GWAS no detectan variantes causales, sino señales de asociación de ciertas regiones. Por lo tanto, se requieren estudios adicionales para poder determinar la variante causal de la región y una aproximación para conseguirlo es la realización de estudios funcionales.

Por medio de herramientas *in silico* como son eQTL (*expression quantitative trait loci*) ha sido posible correlacionar los genotipos de los SNPs asociados previamente con los niveles de expresión de genes localizados en la misma región. Dubois et al. [61] con datos de GWAS realizaron análisis de eQTL en sangre periférica mostrando que 20 de los 38 loci analizados presentan un efecto eQTL significativo, el efecto más significativo se observó en un SNP perteneciente al gen *IL18RAP*. Estos análisis de eQTL se llevaron a cabo en PBMC (células de sangre periférica), pero no hay que excluir genes con una expresión específica a nivel intestinal que desarrollen un papel importante en la enfermedad y que no presenten diferencias significativas en expresión usando células de sangre periférica. [82] Una prueba de ello es el gen *TAGAP*, en el que se observó una diferencia en la base de datos eQTL pero en el estudio de Plaza-Izurieta et al. no se vieron diferencias en los niveles de expresión en tejido duodenal. [83]. Lo mismo ocurre con los genes *SCHIP1* y *LPP* que muestran niveles de expresión alterados en biopsias duodenales [83] y no habían sido previamente asociados con eQTL.[61]

Los resultados de expresión no prueban que el SNP sea la variante causal del locus, pero ayudan a priorizar las líneas de investigación futuras. Existen genes para los que no se conoce su función en la enfermedad y estudios con eQTL y de expresión

en tejidos (por ejemplo biopsias duodenales en ECe) nos pueden ayudar a conocer el mecanismo involucrado en la enfermedad. [74]

## 4. Inmunopatogénesis

Tanto en la EC como en la ECe la respuesta inflamatoria viene caracterizada por una respuesta Th1 en la que el factor de necrosis tumoral α (TNF-α) y el interferón γ (IFN-γ) juegan un papel muy importante. En el caso de la CU se ha visto una implicación de las células Th2. En los últimos años se ha descubierto que en la EII tiene un papel muy importante la respuesta Th17, marcada por la producción de la IL-17.

### 4.1 Inmunopatogénesis de la EII

El conocimiento de la inmunopatogénesis de la EII es todavía parcial. La inflamación crónica parece estar causada por una alteración de la respuesta inmune que implica una activación de células T en respuesta a factores ambientales y/o bacterias comensales del íleon terminal o colon en pacientes susceptibles a padecer la enfermedad. [1, 84]

En la EII el sistema inmunológico puede fallar tanto a nivel de la respuesta innata como adaptativa (figura 2):

Respuesta inmune innata:

La barrera intestinal consiste en una capa epitelial cubierta de mucus donde se localizan numerosos elementos del sistema inmunológico innato (células dendríticas, células de Paneth, macrófagos y neutrófilos). Normalmente la barrera epitelial protege de amenazas bacterianas. En el caso de la EII existe un fallo en la barrera intestinal. Existen defectos en la estructura de los complejos de unión entre células epiteliales, con la disminución de los niveles de expresión de proteínas que forman parte de las *tight juction* en los enterocitos (ocludina, cadherinas y cateninas).[85, 86] En pacientes con EII también se observa una disminución del número de células de Goblet, las

cuales segregan mucinas de carácter protector en el epitelio intestinal. Por otra parte también se ha observado una disminución de las células de Paneth, las cuales se localizan en las criptas intestinales del intestino delgado y secretan péptidos antimicrobianos, y como consecuencia, un descenso de la producción de defensinas.[87-89]

La barrera epitelial representa una primera línea de defensa contra los patógenos, pero también existen otras vías de respuesta innata que actúan a través de la identificación de componentes bacterianos como son los receptores del ámbito extracelular TLR (*Toll-like-receptors*) y receptores del ámbito intracelular NLR (*NOD-like receptors*), en este último se incluye el receptor NOD2 (*nucleotide-binding oligomerization domain containing 2*). Una vez activada la vía TLR y/o NOD2, las células epiteliales producen péptidos antimicrobianos [β-defensina humanas (HBD)]. Tanto en la EC como en CU se ha visto un aumento en la producción de estos péptidos. [89, 90] Algunos autores han postulado que ciertas mutaciones en *NOD2* podrían ocasionar un defecto biológico en las células de Paneth. [87-89, 91] También se ha visto el efecto de ciertas variantes de riesgo en ciertos genes de autofagia, *ATG16L1*, afectando a la funcionalidad de las células de Paneth.[92]

Respuesta adaptativa:

La permeabilidad del epitelio intestinal incrementa el contacto con la flora comensal. Ese exceso de interacciones parece causar una pérdida de tolerancia, activando células dendríticas. En individuos sanos, las células dendríticas presentan antígenos a las células $CD4^+$ favoreciendo su diferenciación a células T reguladoras (Treg) y asegurándose la tolerancia de la flora comensal. En la EII, las células dendríticas activadas por la vía TLR y/o NOD2 producen citocinas proinflamatorias que promueven la diferenciación de las células $CD4^+$ a Th1, Th2 y/o Th17, las cuales conducen a una inflamación persistente.[93] La EC se caracteriza por una elevada producción de IL-2, IL-12, IFNγ y TNFα por parte de las células Th1. [46, 94] Mientras que en la mucosa de los pacientes con CU predomina una infiltración de linfocitos Th2 que se caracterizan por la producción de IL-5, IL-13 y TGFβ. [46, 95]

TNFα es una citocina mediadora entre respuesta innata y adaptativa, su producción se encuentra favorecida tanto por macrófagos activados como por las células Th1. Se trata de una citocina central en la inmunopatogénesis de la EII, ya que

favorece la diferenciación de células estromales a miofibroblastos, que son células promotoras de metaloproteasas de la matriz (MMPs), un tipo de enzima responsable de la degradación tisular. [96]

Estudios recientes hablan de la importancia de una tercera línea celular en la inmunopatogénesis de EII, las Th17. Se trata de células $CD4^+$ $CD25^-$ productoras de IL-17, una citocina proinflamatoria que aparece en la respuesta a bacterias extracelulares, induciendo el reclutamiento de neutrófilos hacia la zona de inflamación [97, 98].

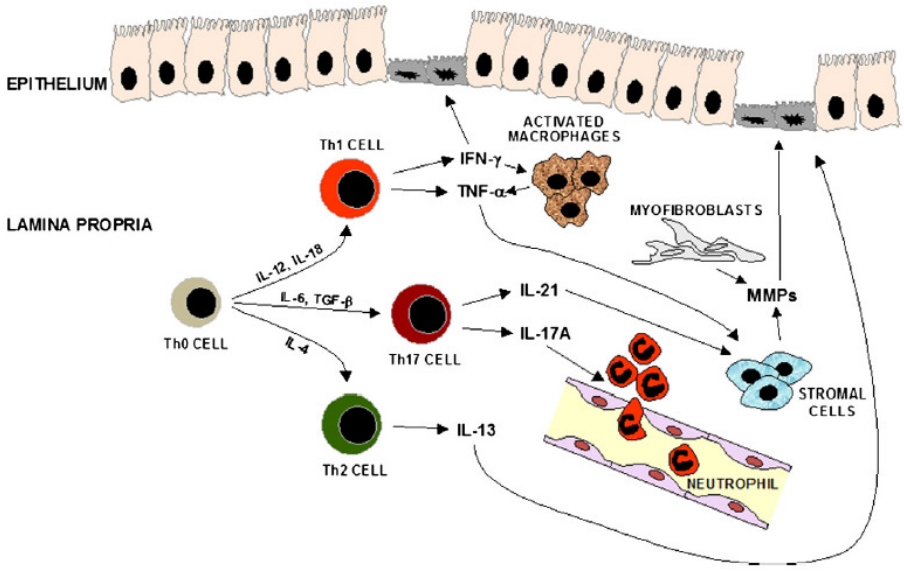



**Figura 2:** Inmunopatogénesis de la enfermedad inflamatoria intestinal. Geremia A. *et al.* [99]

### 4.2 Inmunopatogénesis de la ECe

La clave en la patogénesis de la ECe es el gluten. La predisposición individual hace que tras la ingesta del gluten se activen respuestas innatas y adaptativas con la consiguiente destrucción del epitelio de la mucosa intestinal y la infiltración linfocitaria (figura 3). [100]

Los péptidos de gluten pasan a través de la barrera epitelial del intestino hasta la lámina propria, donde son deamidados por la enzima transglutaminasa tipo 2, una

enzima localizada en el lumen que interviene en la reparación tisular en individuos sanos. Al ser deamidados, los péptidos adquieren carga negativa, lo que les confiere una mayor inmunogenicidad. [101] Son reconocidos en la lámina propria por las células presentadoras de antígenos (APC) y presentados por el HLA de tipo II. [64, 102] Ciertas moléculas de HLA (moléculas DQ2 o DQ8) poseen una mayor afinidad por los péptidos de gluten. [103] Posteriormente las células T $CD4^+$ reconocen al complejo APC-HLA-II y producen citocinas proinflamatorias, principalmente interferón-γ (IFNγ), capaces de desencadenar un conjunto de reacciones que producen finalmente daño en la mucosa intestinal y atrofia de las vellosidades intestinales.[104]También se ha observado que individuos con ECe desarrollan anticuerpos, de gran utilidad para el diagnóstico de la enfermedad, fundamentalmente frente a los péptidos deamidados y la enzima transglutaminasa tisular.[104]

Los linfocitos intraepiteliales (LIEs) están localizados entre las células epiteliales intestinales y la zona basolateral y juegan un papel importante en la inmuno vigilancia del epitelio. Se encuentran involucrados en la inmunopatogénesis de la ECe. Se trata de una población mixta de células T $TCR\alpha\beta^+$, $TCR\gamma\delta^+$ y NK (Natural Killer); aunque la mayoría son células T $CD8^+$ $TCR\alpha\beta^+$. [105] En individuos con ECe activa, el número de LIEs está aumentado, no está claro si se debe a una respuesta al cambio homeostático del medio. LIEs de pacientes con ECe expresan receptores de activación de las células NK (CD94/NKG2C y NKG2D), mientras que en el intestino de individuos sanos expresan receptores inhibitorios de las NK (CD94/ receptor NKG2A). A su vez células del epitelio intestinal de enfermos expresan MIC y HLA-E, ligandos de NKG2D y CD94/NKG2C, respectivamente, aumentando la producción de IFN-γ y produciendo un mayor daño tisular. [106]

Los péptidos de gluten también desarrollan una respuesta innata caracterizada por una expresión elevada de ciertas citocinas como es la IL-15 producida por los enterocitos,[107] citocina que promueve la proliferación y citotoxidad de los linfocitos TCRαβ CD8+.

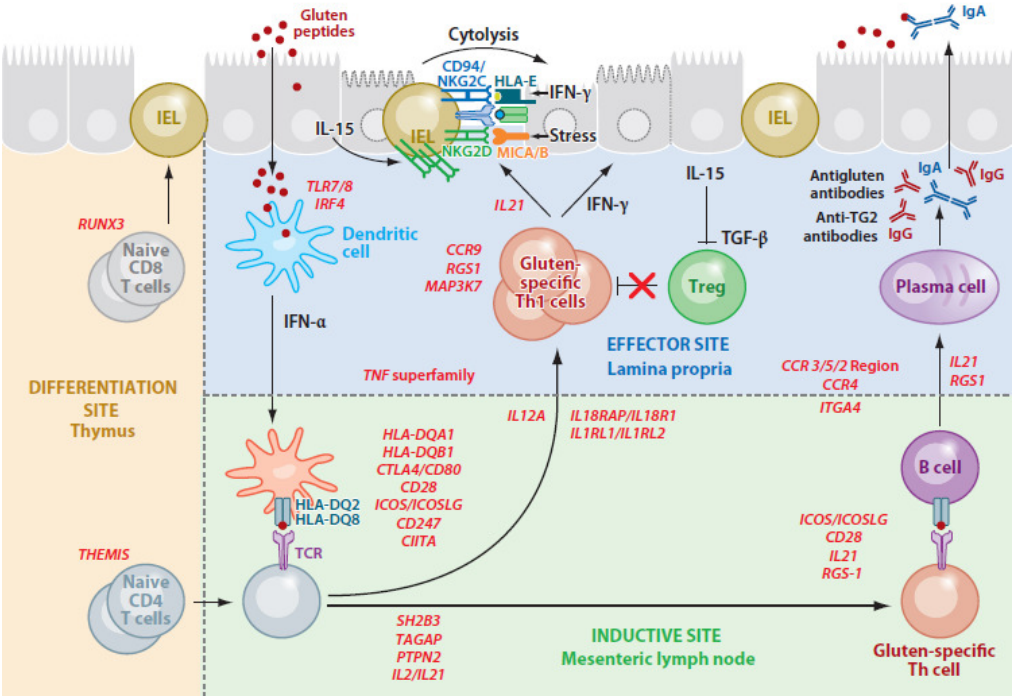



Figura 3. Inmunopatogésis de la Enfermedad Celiaca. Abadie, V *et al.* [78]

### 4.3 Relevancia de las células Th17

Clásicamente EC y ECe se han considerado enfermedades mediadas por una respuesta Th1 y CU tipo Th2. Tras la asociación del gen *IL23R* con el riesgo a desarrollar EC en estudios de barrido genómico, pasó a considerarse de gran importancia el papel de las células Th17. Posteriormente se asociaron a EC otros genes implicados en la respuesta Th17 (*IL12B*, *JAK2*, *STAT3*, *CCR6* y *TNFSF15*). [55, 108, 109]Este tipo de respuesta también se vio asociada a CU [54]y su implicación apoyada por estudios de expresión en biopsias intestinales. [110-112] Estudios en modelos múridos de EII también han mostrado un incremento de los niveles de expresión de IL-12 e IL-23.[111, 113] En ECe no está tan claro este papel, puesto que estudios de expresión muestran niveles alterados de ciertas citocinas relacionadas con las células Th17 en individuos con ECe activa, [114-116] pero en estudios de GWAS no se vieron genes implicados en esta ruta con la excepción de la región IL2/IL21, [59-62] región implicada también en CU y EC y dentro de la cual no se conoce el gen

implicado. IL-21 estimula la proliferación de las células Th17 [117] y regula la inflamación y respuesta de las Th17 en el intestino. [118] Mientras que la IL-2 actúa como un factor de crecimiento de células T y activa a las células Treg. Este subtipo de células T puede suprimir la respuesta efectora de otras células T, de gran importancia para el control de lo propio frente a lo extraño y para prevenir ciertos fallos en el sistema inmunológico que deriven a autoinmunidad. La disminución de actividad o la alteración de la función de las Treg puede producir alteraciones autoinmunes. [119]

La diferenciación de las células T *naive* $CD161^+$ $CD4^+$ a células Th17 requiere la combinación de las citocinas inflamatorias TGF-β, IL-6, IL-23 o IL-1β y de los factores de transcripción ROR-γt y STAT3, que activan la transcripción de IL-17A e IL17-F, que son las principales citocinas efectoras. [120] La producción de IL-21, IL-22 y IL-26 también es característica de este tipo de respuesta, además de ciertos marcadores de superficie: receptor de IL-23 y CCR6 (receptor de quimiocina C-C tipo 6). Las células Th17 parecen estar implicadas en mejorar la protección frente a la infección de bacterias extracelulares y hongos. [120]

La homeostasis intestinal se debe a un balance entre señales anti-inflamatorias y pro-inflamatorias. El sistema inmunológico intestinal tiene que mantener un equilibrio entre las distintas líneas celulares Treg/Th1/Th2/Th17 (figura 4): la tolerancia frente a la flora intestinal y la habilidad para combatir patógenos extracelulares. La inflamación en la EII resulta de una respuesta inadecuada de las células Treg o de una respuesta aumentada por parte de las células Th1/Th17 en el caso de EC y Th2/Th17 en el caso de la CU; existiendo un desequilibrio en el balance del sistema inmunológico intestinal. [121]

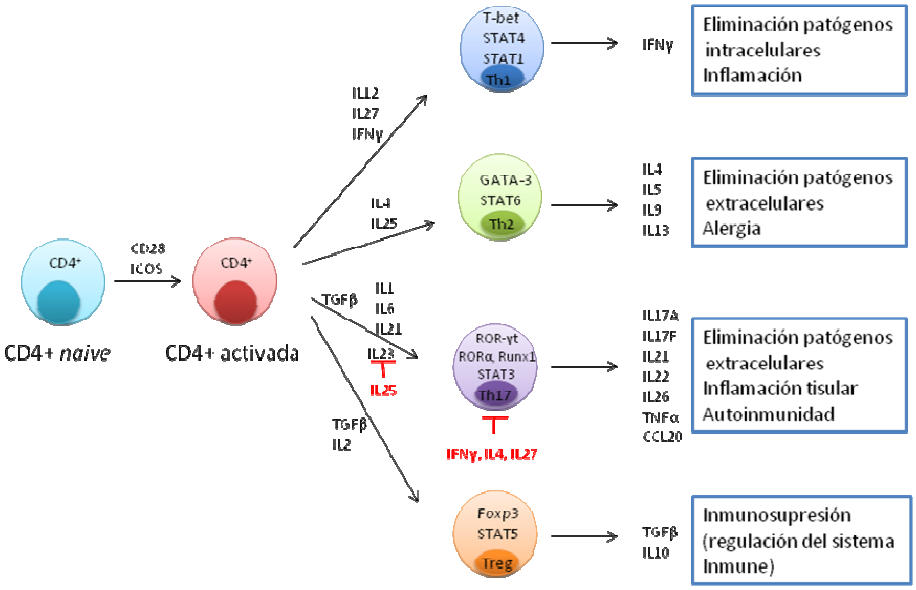



**Figura. 4:** Representación de ciertas líneas celulares implicadas en el sistema inmunológico.

La asociación de los genes implicados en la ruta Th17 en la EII y otras enfermedades abre las puertas al posible estudio de las células Th17 y a la búsqueda de genes implicados en esta ruta como posibles diana terapéutica de tratamientos alternativos en la enfermedad: ciertos anticuerpos monoclonales que actúan a nivel de la ruta de las células Th17 se encuentran en fase de estudio.[122-124]

## 5. Tratamiento

El tratamiento en la ECe es la eliminación de gluten en la dieta. Tras la retirada del gluten la respuesta en la mayoría de los pacientes es efectiva aunque los títulos de anticuerpos no desaparecen hasta meses o incluso años. [104] Pacientes con un tratamiento estricto de dieta sin gluten reducen el riesgo de desarrollar otras enfermedades autoinmunes [125] y linfoma T asociado a enteropatía. [126, 127]

Existe un grupo minoritario (alrededor del 5%) de pacientes refractarios (ECeR), que no responden adecuadamente a la retirada del gluten, presentando un mayor riesgo de desarrollar complicaciones, [128] de ahí la necesidad de buscar alternativas terapéuticas.

Debido al papel etiológico del gluten, la ECe es un buen modelo de estudio. La diferencia en los niveles de expresión de ciertos genes en biopsias intestinales de enfermos tanto en dieta libre de gluten como sin ella puede dar una idea de cual es la ruta involucrada como consecuencia o causa de la enfermedad.[83]

En el caso de la EC la variedad de tratamientos es amplia, desde corticoides hasta tratamientos biológicos como el infliximab (IFX). El IFX es un anticuerpo monoclonal IgG1 neutralizante del TNFα, es tratamiento de elección para pacientes con EC refractaria que no responden a una terapia convencional con tratamiento anti-inflamatorios o inmunomoduladores.[129] Se ha visto una buena respuesta a este fármaco en pacientes con afectación luminal y/o fistulizante. [130] A pesar de ser un tratamiento de gran eficacia, existe un 20-30% de pacientes que no responden a este fármaco, de ahí la necesidad de encontrar ciertos marcadores que ayuden a predecir cuál va a ser la respuesta y promover una medicina individualizada.

Se han apuntado diversos factores que parecen influir en la respuesta a IFX: años de duración, fumar, [130-133] administración de tratamiento concomitante [134, 135] y fenotipo de la enfermedad.[130-133] Además el perfil genético de cada individuo también influye en este tipo de respuesta. La búsqueda de genes que modulan la respuesta a los fármacos es el objetivo de la farmacogenética. En relación a IFX los estudios farmacogenéticos se han centrado en genes con un papel importante en la inmunopatogénesis de la enfermedad como es el *NOD2* [136, 137] y genes relacionados con el TNFα [138], sin llegar a encontrar ninguna asociación confirmada con la respuesta. Por tanto, se necesitan más estudios para optimizar el uso de IFX y ayudar a conocer las vías de acción de éste fármaco. La combinación de asociaciones genéticas y los perfiles de expresión en tejido podría ayudar a buscar nuevas dianas para el desarrollo de tratamientos anti-inflamatorios.

En ambas enfermedades es importante un diagnóstico rápido de la enfermedad, ya que con la aplicación de tratamiento se reducen el número de complicaciones. [139-143]

- Búsqueda de factores genéticos de susceptibilidad a enfermedad celiaca o enfermedad inflamatoria intestinal que ayuden a explicar la "heredabilidad perdida":

  - Búsqueda de nuevos factores de susceptibilidad a enfermedad celiaca en la región del HLA.

  - Confirmación de interacciones gen-gen (*IL23R*, *CARD15* y *TLR9*) como factores de riesgo a enfermedad inflamatoria intestinal.

  - Estudio de la asociación de genes que intervienen en la ruta Th17 con la susceptibilidad a padecer enfermedad celiaca.

- Estudio de la expresión a nivel intestinal de genes implicados en la ruta Th17 en enfermedad celiaca y en colitis ulcerosa.

- Estudio de factores genéticos que ayuden a predecir la respuesta a infliximab en pacientes con enfermedad de Crohn:

  - Estudio de *TNFRSF1A* y *TNFRSF1B*, previamente asociados con la respuesta a infliximab en población japonesa.

  - Estudio de genes cuya expresión conjunta parece predecir la respuesta a infliximab en pacientes de Crohn colónico: *TNFAIP6*, *IL11*, *G0S2*, *S100A8* y *S100A9*.

1.- ESTUDIO DE NUEVOS FACTORES GENÉTICOS EN LA BÚSQUEDA DE LA "HEREDABILIDAD PERDIDA"

***I.*** ***HLA and celiac disease susceptibility: new genetic factors bring open questions about the HLA influence and gene-dosage effects.***

***II.*** ***Interaction between TLR9 and IL23R polymorphisms influences Crohn´s disease susceptibility.***

***III.*** ***Th17-related genes and celiac disease susceptibility.***

**I.** ***HLA and celiac disease susceptibility: new genetic factors bring open questions about the HLA influence and gene-dosage effects.***

Luz María Medrano[1], MSc, Bárbara Dema[1], PhD, Arturo López-Larios[1], MD, Carlos Maluenda[2], MD and PhD, Andrés Bodas[2], MD and PhD, Natalia López-Palacios[3], MD and PhD, M. Ángeles Figueredo[1], MD and PhD, Miguel Fernández-Arquero[1], PhD and Concepción Núñez[1], PhD

1 UGC de Inmunología, Hospital Clínico San Carlos, Instituto de Investigación Sanitaria del Hospital Clínico San Carlos (IdISSC), Madrid, Spain
2 Servicio de Pediatría, Hospital Clínico San Carlos, Instituto de Investigación Sanitaria del Hospital Clínico San Carlos (IdISSC), Madrid, Spain
3 Servicio de Aparato Digestivo, Hospital Clínico San Carlos, Instituto de Investigación Sanitaria del Hospital Clínico San Carlos (IdISSC), Madrid, Spain

**PLoS One. 2012;7(10):e48403**

**Abstract**

Celiac disease (CD) is a chronic inflammatory disorder triggered after gluten ingestion in genetically susceptible individuals. The major genetic determinants are *HLA-DQA1*05* and *HLA-DQB1*02*, which encode the DQ2 heterodimer. These alleles are commonly inherited in cis with *DRB1*03:01*, which is associated to numerous immune-related disorders, in some cases contributing with different risk depending on the haplotype context. We aimed at investigating those possible differences involving *DRB1*03:01*-carrying haplotypes in CD susceptibility. A family (274 trios) and a case-control sample (369 CD cases/461 controls) were analyzed. *DRB1*03:01*-carrying individuals were classified according to the haplotype present (ancestral haplotype (AH) 8.1, AH 18.2 or non-conserved haplotype) after genotyping of *HLA-DRB1*, *-DQA1*, *-DQB1*, *-B8, TNF* -308, *TNF* -376 and the *TNFa* and *TNFb* microsatellites. We observe that the AH 8.1 confers higher risk than the remaining *DRB1*03:01*-carrying haplotypes, and this effect only involves individuals possessing a single copy of *DQB1*02*. CD risk for these individuals is similar to the one conferred by inherit *DQA1*05* and *DQB1*02* in trans. It seems that an additional CD susceptibility factor is present in the AH 8.1 but not in other *DRB1*03:01*-carrying haplotypes. This factor could be shared with individuals possessing DQ2.5 trans, according to the similar risk observed in those two groups of individuals.

**Key words:** ancestral haplotype 8.1, celiac disease susceptibility, *DRB1*03:01*, DQ2.5 cis, DQ2.5 trans

**Introduction**

Human leukocyte antigen (HLA) is a master piece in the pathogenesis of celiac disease (CD), as first evidenced by the strong genetic association existent between CD susceptibility and certain HLA alleles. This region, located on 6p21, contains hundreds of genes with immunological function and it is characterized by a high gene density and variability and an extensive linkage disequilibrium, which make difficult to pinpoint the causal variant/s. Despite this, CD can be considered a particular disease since the specific HLA alleles involved and their functional implication are well-established. The presence of the *DQA1*05* and *DQB1*02* susceptibility alleles implies the formation of the α and β chains of the HLA-DQ2 heterodimer, present in around 90-95% of CD individuals. This molecule shows high affinity for peptides resultant from incomplete gluten digestion, which bind and present to antigen specific T cells, triggering the intestinal inflammation prototypical of CD.

In most cases, *DQA1*05* and *DQB1*02* are encoded in the same chromosome (DQ2.5 cis) and appear in very strong linkage disequilibrium with *DRB1*03:01*. In fact, this allele was first associated with CD risk [150]. The *DQA1*05*, *DQB1*02* and *DRB1*03:01* alleles can be present in two different haplo-specific contexts and constitute the so-called ancestral haplotypes (AH) 8.1 and AH 18.2; they can also be found within non-specific allelic combinations and constitute other less frequent haplotypes (hereafter called non-conserved haplotypes). The *DRB1*03:01* allele, and consequently, *DRB1*03:01* haplotypes, have been associated to numerous immune-mediated disorders, as type 1 diabetes, multiple sclerosis or selective IgA deficiency, among many others; in some cases with the different *DRB1*03:01* haplotypes showing a differential contribution to disease risk [72, 73]. A differential behaviour between AH 18.2 and AH 8.1 has also been described in CD [151, 152], although no relevance has been given to this observation. The *DQA1*05* and *DQB1*02* alleles can also be inherited in trans, each encoded in one chromosome from each parent (DQ2.5 trans).

HLA influence on CD susceptibility shows a dose effect. Individuals can be classified in high or intermediate CD risk according to the number of *DQA1*05*- and *DQB1*02*-carrying alleles. Homozigosity for DQ2.5 cis and heterozigosity for DQ2.5 cis with a chromosome possessing a second *DQB1*02* allele (DQ2.2) confer the

highest risk to develop CD. Heterozigosity for DQ2.5 cis in individuals with a single copy of *DQB1*02* (non-DQ2.2) or presence of DQ2.5 trans confer intermediate risk.

Additionally to the molecule DQ2.5, the influence of HLA-DQ8 (genetically *DQA1*03*, *DQB1*03:02*) on the disease is already known. This molecule is present in almost all the CD patients without DQ2.5. However, the genetic influence of the HLA region in CD is not limited to the factors coding DQ2 or DQ8, and several works have attempted to discover new susceptibility factors without much success (see [153] for review). Some variants in the *TNF* gene have been suggested as DQ2 independent factors for CD susceptibility, even as the responsible factors for the additional risk present on the AH 8.1. [154]. Last years have witnessed a spectacular increase in the knowledge of the genetic basis of CD, favoured by development of genome wide association studies (GWAS), but these works have not added new information about the HLA contribution because they have been mainly focused on the influence of genes outside this region.

We aimed at investigating the additional genetic contribution to CD susceptibility lying on the HLA region, by focusing in the possible differential contribution of the different *DRB1*03:01*-carrying haplotypes.

## Materials and Methods

### Ethics Statement

This study was approved by the ethical committee (CEIC) of Hospital Clínico San Carlos. Samples were obtained after obtaining written informed consent. For children, the informed consent was signed by their parents or legal guardian.

### Subjects

A total of 274 trios composed for both parents and the affected child and a case-control series consisting of 369 independent CD patients and 461 ethnically matched healthy controls were studied. CD patients were diagnosed following the European Society for Pediatric Gastroenterology, Hepatology, and Nutrition (ESPGHAN) [71], 97% are positive for HLA-DQ2 and/or HLA-DQ8. Controls correspond mainly to blood donors and laboratory staff. CD samples were consecutively collected in two centres of the same region (Hospital La Paz and Hospital Clínico San Carlos, Madrid,

Spain) and controls were collected at the Hospital Clínico San Carlos. All samples correspond to unrelated Spanish white individuals.

**Genotyping**

DNA was extracted from fresh peripheral blood leukocytes by a "salting out" procedure. All samples were genotyped for *HLA-DRB1*, *-DQA1* and *-DQB1* by PCR-SSOP (Polymerase Chain Reaction-Sequence Specific Oligonucleotide Probe). The different *DRB1*03:01* haplotypes were assessed by additional genotyping of the *TNF* single nucleotide polymorphisms (SNPs) -308 (rs1800629) and -376 (rs1800750) and the microsatellites *TNFa* and *TNFb*; those polymorphisms were typed as previously described [155, 156]. The presence of the *HLA-B8* allele was tested by TaqMan technology using the tag SNPs rs6457374 and rs2844535 (Applied Biosystems Inc., Foster City, CA, USA).

We defined the AH 8.1 according to the simultaneously presence of *DRB1*03:01, DQA1*05:01, DQB1*02:01, TNF -308A, TNFa2b3* and *B*8*. AH 18.2 was defined according to carriage of *DRB1*03:01, DQA1*05:01, DQB1*02:01, TNF -376A* and *TNFa1b5*. Haplotypes with all the remaining allelic combinations in those loci or markers were designed as non-conserved haplotypes. The *TNF* markers studied were selected because, in *DRB1*03:01* subjects, they are haplo-specific for AH 18.2 or AH 8.1.

**Statistical analysis**

HLA haplotypes were deduced directly from the pedigree for patients used in the family study. In cases and controls, the EM (Expectation-Maximization) algorithm implemented in the Arlequin software was used to estimate haplotype frequencies.

The transmission disequilibrium test (TDT) was used to analyse the preferential transmission of one haplotype over the others when analysing family data. This test uses only information provided by heterozygous parents.

Comparisons between groups were performed with the chi-square test using the statistical package EpiInfo v5.00 (CDC, Atlanta, USA). Heterogeneity between haplotype groups was evaluated with Review Manager (RevMan) 5.0 software (Copenhagen: The Nordic Cochrane Centre, The Cochrane Collaboration, 2008).

**Results**

We studied 274 trios to investigate the possibility of a differential transmission of the different *DRB1*03:01* haplotypes to the affected child (Table 1). *DRB1*03:01* is always preferentially transmitted, independently of its haplotype context (see TDT results, in Table 1). However, the distortion in the transmission of this allele is significantly higher when it is present in the AH 8.1, compared with its presence in the remaining *DRB1*03:01*-containing haplotypes ($p=8.7*10^{-4}$ vs. AH 18.2; $p=2.4*10^{-4}$ vs. non-conserved haplotypes). AH 18.2 and non-conserved haplotypes show a similar preferential transmission to offspring (p=0.99). These differences are also observed when considering the haplotype transmission from the *DRB1*03:01* homozygous parents (composed by different *DRB1*03:01* haplotypes) included in the 274 families (Table 2).

**Table 1.** Transmission data of the diverse *DRB1*03:01*-containing haplotypes in the 274 families studied and TDT results

| ***DRB1*03:01* HAPLOTYPES** | **T** | **U** | **TDT** |
|---|---|---|---|
| 1. All | 268 | 88 | $p<10^{-10}$ |
| 2. AH 8.1 | 118 | 18 | $p<10^{-10}$ |
| 3. AH 18.2 | 58 | 27 | $p= 2.2*10^{-4}$ |
| 4. Non-conserved | 92 | 41 | $p=4.9*10^{-6}$ |
| 5. Non-AH 8.1 | 150 | 68 | $p=5.3*10^{-9}$ |

T=transmitted; U=untransmitted; TDT=transmission disequilibrium test.
Non-AH 8.1 includes AH18.2 and non-conserved haplotypes.
TDT results are calculated after excluding homozygous parents (5 for AH 8.1, 5 for AH 18.2 and 7 for *DRB1*03:01* non-conserved haplotypes).
Haplotype comparisons: 2 vs. 3: p=0.00087 OR=3.05 95% CI 1.48-6.33; 2 vs. 4: p=0.00024 OR=3.06 95% CI 1.59-5.94; 3 vs. 4: p=0.99 OR=1.00 95% CI 0.54-1.88; 2 vs. 5: $p=7.8*10^{-5}$ OR=3.06 95% CI 1.67-5.65.

**Table 2.** Transmission data of the diverse *DRB1*03:01*-containing haplotypes from *DRB1*03:01* homozygous parents (N=27)

| HAPLOTYPE COMPOSITION | N | T | TDT |
|---|---|---|---|
| AH 8.1 - AH 18.2 | 6 | 6 AH 8.1, 0 AH 18.2 | p=0.016 |
| AH 8.1 - Non-conserved | 10 | 7 AH 8.1, 3 Non-conserved | p=0.17 |
| AH 18.2 - Non-conserved | 11 | 4 AH 18.2, 7 Non-conserved | p=0.27 |
| AH 8.1 – Non-AH 8.1 | 16 | 13 AH 8.1, 3 Non- AH 8.1 | p=0.011 |

T=transmitted; TDT=transmission disequilibrium test.

We wanted to validate this observation in a case-control sample (Table 3). Since no differences were observed between AH 18.2 and non-conserved haplotypes in the family data, they were combined in subsequent analysis (and called non-AH 8.1). As already known, *DRB1*03:01* overall appears at significantly higher frequency in CD patients than in controls: 45% vs. 14%, respectively (OR=4.97 95% CI 3.90-6.34, $p<10^{-7}$); but we additionally show that this case-control difference is higher when considering only the AH 8.1 (OR=6.53 95% CI 4.47-9.56, $p<10^{-7}$).

**Table 3.** Frequency and comparison of the diverse *DRB1*03:01* haplotypes in celiac disease (CD) patients (2N=738) and controls (2N=922), overall and classified according to the different HLA CD risk categories

| ***DRB1*03:01* HAPLOTYPES** | **CD** | | **CONTROLS** | | **CD vs. CONTROLS** | | **HETEROGENEITY AH 8.1 VS. NON-AH 8.1** | |
|---|---|---|---|---|---|---|---|---|
| | **2N** | **%** | **2N** | **%** | **p** | **OR (95% CI)** | **p** | **$I^2$** |
| **ALL *DRB1*03:01*** | | | | | | | | |
| All | 330 | 44.7 | 129 | 14.0 | $<10^{-7}$ | 4.97 (3.90-6.34) | | |
| AH 8.1 | 141 | 19.1 | 42 | 4.6 | $<10^{-7}$ | 6.53 (4.47-9.56) | 0.06 | 71% |
| Non-AH 8.1 | 189 | 25.6 | 87 | 9.4 | $<10^{-7}$ | 4.22 (3.16-5.65) | | |
| **DQ2.5 cis+DQ2.5 cis** | | | | | | | | |
| All | 74 | 10.0 | 16 | 1.7 | $<10^{-7}$ | 8.99 (5.03-16.28) | | |
| AH 8.1 | 30 | 4.1 | 6 | 0.65 | $<10^{-7}$ | 9.72 (3.93-28.74) | 0.82 | 0% |
| Non- AH 8.1 | 44 | 6.0 | 10 | 1.1 | $<10^{-7}$ | 8.55 (4.10-18.31) | | |
| **DQ2.5 cis+DQ2.2** | | | | | | | | |
| All | 104 | 14.1 | 13 | 1.4 | $<10^{-7}$ | 15.55 (8.39-29.37) | | |
| AH 8.1 | 40 | 5.4 | 6 | 0.7 | $<10^{-7}$ | 12.96 (5.39-37.64) | 0.60 | 0% |
| Non- AH 8.1 | 64 | 8.7 | 7 | 0.8 | $<10^{-7}$ | 17.77 (8.03-46.30) | | |
| **DQ2.5cis+non-DQ2.2** | | | | | | | | |
| All | 152 | 20.6 | 100 | 10.8 | $<10^{-7}$ | 2.95 (2.21-3.94) | | |
| AH 8.1 | 72 | 9.8 | 30 | 3.3 | $<10^{-7}$ | 4.66 (2.94-7.44) | 0.009 | 85% |
| Non-AH 8.1 | 80 | 10.8 | 70 | 7.6 | $3.2*10^{-6}$ | 2.22 (1.56-3.17) | | |

% are referred to the total number of haplotypes.

Non-AH 8.1 includes AH 18.2 and non-conserved haplotypes.

DQ2.5 cis= *DQB1*02:01-DQA1*05:01*; DQ2.2= *DQB1*02:02-DQA1*02:01*.

In CD, it is well established the existence of a dose effect, what implies the existence of different CD risk categories attending to their HLA constitution. We investigated this differential risk contribution of *DRB1*03:01*-containing haplotypes in those categories and found that only in those individuals carrying a single copy of the *DQB1*02* allele (individuals DQ2.5 cis + non-DQ2.2 in Table 3), the presence of AH 8.1 confers additional risk.

When considering the HLA risk categories according to gene dosage effects, carriage of the DQ2 molecule in individuals with a single copy of the *DQB1*02* allele is considered as conferring intermediate CD risk. A similar risk is conferred by the presence of DQ2.5 trans, although some groups reported this to be an intermediate higher risk group [157]. We compared CD risk in carriers of DQ2.5 trans (48 individuals out of 369 patients and 21 out of 461 controls) to CD risk in carriers of DQ2.5 cis with and without the AH 8.1 and we observed that DQ2.5 trans confers similar risk than DQ2.5 cis with AH 8.1 (heterogeneity: p=0.91, $I^2$=0%) and significantly higher risk than DQ2.5 cis with non-AH 8.1 haplotypes (heterogeneity: p=0.09, $I^2$=64%).

Finally, we investigated the possibility that the similar risk conferred by DQ2.5 trans and DQ2.5 cis with AH 8.1 was due to the presence of a common susceptibility factor. In most cases, carriers of the molecule DQ2.5 trans are genetically characterized by being heterozygous *DQB1*03:01-DQA1*05:05/DQB1*02:02-DQA1*02:01* (serologically DR5/DR7, terms also used hereafter for simplification purposes). We used genotype data corresponding to 6,769 SNPs located in the HLA (29.96-33.19 Mb interval), which were previously obtained in a subset of our Spanish samples (more than 500 CD patients and 300 controls) in the context of the Immunochip Project (http://www.immunobase.org). We selected all the homozygous individuals for AH 8.1, AH 18.2, DR5 (*DQB1*03:01-DQA1*05:05*) and DR7 (*DQB1*02:02-DQA1*02:01*) and looked for a chromosomal region shared between AH 8.1 and DR5 (*DQB1*03:01-DQA1*05:05*) or AH 8.1 and DR7 (*DQB1*02:02-DQA1*02:01*) but not common to AH 18.2. However, no regions emerged after that search. Note that this result could be affected by the low number of available homozygous individuals, mainly DR5 (*DQB1*03:01-DQA1*05:05*) and DR7 (*DQB1*02:02-DQA1*02:01*) (3 and 4, respectively).

**Discussion**

Our results evidence the presence of an additional susceptibility factor to CD in the HLA region, which is linked to AH 8.1. This factor only increases susceptibility when appearing in individuals carrying a single copy of the *DQB1*02* allele (DQ2.5 cis non-DQ2.2) and split up the HLA intermediate risk group into two groups: one with higher intermediate risk, which is composed of carriers of the molecule DQ2.5 cis encoded by AH 8.1 and carriers of the molecule DQ2.5 trans; and one with lower intermediate risk, which is composed of carriers of the molecule DQ2.5 cis encoded by either AH 18.2 or non-conserved *DRB1*03:01* haplotypes.

The well-known contribution of the DQ2 molecule to CD pathogenesis is genetically based on carriage of *DQA1*05* and *DQB1*02*. These two alleles are commonly present in the small segment identical by descent among *DRB1*03:01*-containing haplotypes. However, outside that shared segment, the divergence between *DRB1*03:01* haplotypes do not differ from that found between disparate haplotypes [158]; therefore, a susceptibility factor located there is not expected to be present in all *DRB1*03:01* haplotypes. On the other hand, previous studies suggested a close evolutionary relationship among *DRB1*03:01*-containing haplotypes, DR5 and DR7 (*DQB1*03:01-DQA1*05:05* and *DQB1*02:02-DQA1*02:01*) [159], which could explain the existence of a common susceptibility factor between AH 8.1 and one of those haplotypes. No definitive conclusion can de drawn from our data due to the low number of chromosomes compared. Further analysis including higher sample size is mandatory.

The analysis of the HLA region in GWAS showed, besides the expected peak corresponding to DQ2.5 cis, two SNPs associated to CD risk, both located within or adjacent to *HLA-DQA1* and *HLA-DQB1* [58]. This is in accordance with our results, which suggest a risk factor common to *DRB1*03:01*-containing haplotypes and DR5 or DR7 (*DQB1*03:01-DQA1*05:05* or *DQB1*02:02-DQA1*02:01*), because those are the regions that they shared.

One intriguing issue emerging from our study is why the additional risk factor present in the AH 8.1 does not seem to influence on CD susceptibility when it appears in individuals carrying a second copy of the *DQB1*02* allele. In CD, T cell stimulation due to gluten-derived peptides depends on the number and type of HLA-DQ2 molecules expressed. DQ2.5 molecules can bind a high repertoire of gluten peptides, but only a restricted subset is bound to DQ2.2 molecules, which reduce the

immunogenicity of DQ2.2. Additionally, the number of these DQ molecules is also a relevant factor in T cell stimulation and this depends on the number of specific alleles in *DQA1* and *DQB1* loci, which determines the possible αβ-chain combinations constituting the DQ heterodimers [160]. As a matter of fact, all HLA-DQ molecules are identical in HLA-DQ2.5 homozygous individuals, which can bind a very high repertoire of gluten peptides and confer the highest CD risk. It could be speculated that in such scenario an additional susceptibility factor has null or limited possibility to increase risk. By the contrary, in HLA-DQ2.5 cis individuals (without a second copy of *DQB1*02, i.e.*, non-DQ2.2), only one of the four possible αβ combinations constitutes an HLA-DQ2.5 molecule and therefore the presence of a genetic factor which increases immunogenicity against gluten derived peptides could have a relevant impact in increasing CD risk.

The HLA dose effect is also influenced by differences in the kinetic stability of the interaction between HLA molecules and gluten derived peptides, key factor for development of T cell responses against gluten [68, 161]. For most peptide ligands, DQ2.5 shows higher binding stability than DQ2.2. The risk variant present in the AH 8.1, and presumably in individuals possessing DQ2.5 trans molecules, could affect the kinetic stability of the interaction HLA-gluten either increasing the number of gluten derived peptides which can bind with high affinity or increasing the binding stability of peptides previously recognised. Bodd *et al* [161] claimed that T-cell epitopes must be assessed and characterized in the context of the HLA molecules expressed by the T-cell donor and underlined the relevance that this could have for future peptide-based vaccines. According to that, it would be interesting to establish a comparison of the HLA-DQ molecules present in individuals DQ2.5 cis with AH 8.1, DQ2.5 cis with non-AH 8.1 haplotypes and DQ2.5 trans. Differences in their binding gluten peptides would imply that peptide-based vaccines should look at those individuals differentially.

Much more work deserves this field, with several open questions as which is the causal variant lying on the AH 8.1 and which are their specific functional implications.

**Acknowledgements**

We are most grateful to Carmen Martínez Cuervo and M. Ángel García Martínez for their expert technical assistance.

**References**

1. Keuning JJ, Pena AS, van Leeuwen A, van Hooff JP, va Rood JJ (1976) HLA-DW3 associated with coeliac disease. Lancet 1: 506-508.

2. Baschal EE, Aly TA, Jasinski JM, Steck AK, Noble JA, et al. (2009) Defining multiple common "completely" conserved major histocompatibility complex SNP haplotypes. Clin Immunol 132: 203-214.

3. de la Concha EG, Cavanillas ML, Cenit MC, Urcelay E, Arroyo R, et al. (2012) DRB1*03:01 Haplotypes: Differential Contribution to Multiple Sclerosis Risk and Specific Association with the Presence of Intrathecal IgM Bands. PLoS One 7: e31018.

4. Bolognesi E, Karell K, Percopo S, Coto I, Greco L, et al. (2003) Additional factor in some HLA DR3/DQ2 haplotypes confers a fourfold increased genetic risk of celiac disease. Tissue Antigens 61: 308-316.

5. Bilbao JR, Calvo B, Aransay AM, Martin-Pagola A, Perez de Nanclares G, et al. (2006) Conserved extended haplotypes discriminate HLA-DR3-homozygous Basque patients with type 1 diabetes mellitus and celiac disease. Genes Immun 7: 550-554.

6. Louka AS, Lie BA, Talseth B, Ascher H, Ek J, et al. (2003) Coeliac disease patients carry conserved HLA-DR3-DQ2 haplotypes revealed by association of TNF alleles. Immunogenetics 55: 339-343.

7. de la Concha EG, Fernandez-Arquero M, Vigil P, Rubio A, Maluenda C, et al. (2000) Celiac disease and TNF promoter polymorphisms. Hum Immunol 61: 513-517.

8. Husby S, Koletzko S, Korponay-Szabo IR, Mearin ML, Phillips A, et al. (2012) European Society for Pediatric Gastroenterology, Hepatology, and Nutrition guidelines for the diagnosis of coeliac disease. J Pediatr Gastroenterol Nutr 54: 136-160.

9. Nedospasov SA, Udalova IA, Kuprash DV, Turetskaya RL (1991) DNA sequence polymorphism at the human tumor necrosis factor (TNF) locus. Numerous TNF/lymphotoxin alleles tagged by two closely linked microsatellites in the upstream region of the lymphotoxin (TNF-beta) gene. J Immunol 147: 1053-1059.

10. Fernandez-Arquero M, Arroyo R, Rubio A, Martin C, Vigil P, et al. (1999) Primary association of a TNF gene polymorphism with susceptibility to multiple sclerosis. Neurology 53: 1361-1363.

11. van Heel DA, Hunt K, Greco L, Wijmenga C (2005) Genetics in coeliac disease. Best Pract Res Clin Gastroenterol 19: 323-339.

12. Traherne JA, Horton R, Roberts AN, Miretti MM, Hurles ME, et al. (2006) Genetic analysis of completely sequenced disease-associated MHC haplotypes identifies shuffling of segments in recent human history. PLoS Genet 2: e9.
13. Sollid LM (2000) Molecular basis of celiac disease. Annu Rev Immunol 18: 53-81.
14. van Heel DA, Franke L, Hunt KA, Gwilliam R, Zhernakova A, et al. (2007) A genome-wide association study for celiac disease identifies risk variants in the region harboring IL2 and IL21. Nat Genet 39: 827-829.
15. Vader W, Stepniak D, Kooy Y, Mearin L, Thompson A, et al. (2003) The HLA-DQ2 gene dose effect in celiac disease is directly related to the magnitude and breadth of gluten-specific T cell responses. Proc Natl Acad Sci U S A 100: 12390-12395.
16. Fallang LE, Bergseng E, Hotta K, Berg-Larsen A, Kim CY, et al. (2009) Differences in the risk of celiac disease associated with HLA-DQ2.5 or HLA-DQ2.2 are related to sustained gluten antigen presentation. Nat Immunol 10: 1096-1101.
17. Bodd M, Kim CY, Lundin KE, Sollid LM (2012) T-Cell Response to Gluten in Patients With HLA-DQ2.2 Reveals Requirement of Peptide-MHC Stability in Celiac Disease. Gastroenterology 142: 552-561.

**II.** ***Interaction between TLR9 and IL23R polymorphisms influences Crohn´s disease susceptibility***

Luz María Medrano[1], Bárbara Dema[1], Juan Luis Mendoza[2], Carlos Taxonera[2], Manuel Díaz-Rubio[2], Emilio G de la Concha[1], Elena Urcelay[1] and Concepción Núñez[1]

[1]Clinical Inmunology Department, Hospital Clínico San Carlos, Madrid, Spain
[2] Inflammatory Bowel Disease Unit, Hospital Clínico San Carlos, Madrid, Spain

**(Enviado)**

**Abstract**

OBJECTIVES: Crohn´s disease (CD) shows a complex etiology with multiple genes involved. Genetic interactions between polymorphisms in *TLR9* with variants in *CARD15* and *IL23R* have been described as modulating CD susceptibility in the German population. We aimed at replicating those interactions in our independent CD samples from the Spanish population and at corroborating the statistically significant interactions in ulcerative colitis (UC) patients.

METHODS: Nine single nucleotide polymorphisms (SNPs) located in the *CARD15* (rs2066844=R702W, rs2066845=G908R and rs2066847=1007fsinsC), *IL23R* (rs1004819, rs7517847 and rs11209026) and *TLR9* (rs352162, rs187084 and rs5743836) genes were genotyped in 416 CD patients and 547 unrelated healthy controls. Some of those SNPs were also genotyped for replication purposes in an additional set of 122 CD patients, 307 controls and 452 UC patients. Genetic interactions were evaluated using chi-square tests in *TLR9/CARD15* and *TLR9/IL23R* stratified tables.

RESULTS: A significant interaction between *TLR9* rs352162 and *IL23R* rs7517847 was observed and replicated in CD patients but not in controls or UC patients. Carriage of the minor allele of *IL23R* rs7517847 confers a strong protective effect towards CD development when present in *TLR9* rs352162_CC homozygous individuals: p=0.0003 OR=0.44 95% CI 0.27-0.71. Genetic interactions were not observed between *TLR9* and *CARD15* variants.

CONCLUSIONS: We confirmed the presence of epistasis between *TLR9* and *IL23R* polymorphisms affecting CD susceptibility in the Spanish population.

**Key words:** inflammatory bowel disease susceptibility, epistasis, toll-like receptor-9, interleukin-23 receptor**,** *CARD15*

**Introduction**

Crohn´s disease (CD) and ulcerative colitis (UC), the two main forms of inflammatory bowel disease (IBD), show a complex etiology with several genetic and environmental factors involved. Several aspects distinguish those two diseases although both are characterized by chronic inflammation of the mucosal surface of the gastrointestinal tract most probably resulting from an inappropriate response to intestinal microbes. Some genes involved are specific but others are common to both clinical entities.

In last years, genome wide association studies unravelled many genes involved in these and other complex diseases, which highly contributed to expand our knowledge about possible functional processes which could be involved in those diseases. Specifically, the studies developed in relation to CD increase in more than 30 the number of susceptibility loci previously related with this disease [125]. However, despite this huge advance, a high proportion of heritability remains to be explained. Different approaches have been focused on diminishing the proportion of that missing heritability and, although great progresses have not been achieved, genetic interactions studies have occasionally reported interesting results [162-165]. In CD, epistasis between polymorphisms in the *TLR9* gene with CD-associated variants in *CARD15* and *IL23R* has been previously described in the German population [166]. The functional implications of these genes give more relevance to these interactions. TLR-9 and CARD15 are both related with recognition of pathogen associated molecular patterns and the subsequent development of an innate immune response: CARD15 recognizes muramyl dipeptide present in bacterial cell walls, and TLR-9 participates in recognition of unmethylated CpG dinucleotides present in bacterial DNA. IL-23R is mainly related with a Th-17 immune response, which could be stimulated by bacterial components trough pattern recognition receptors.

In this work, we tried to replicate the genetic interaction between polymorphisms in *TLR9* with variants in *CARD15* and *IL23R* to validate the results in an independent population. The significant results obtained were subsequently tested using UC patients, based on the partially common genetic background between the two IBD forms.

## Material and Methods

### Subjects

We studied a sample comprising a maximum of 416 CD patients, 452 UC patients and 547 ethnically matched unrelated controls. For replication purposes, an additional sample of 122 CD patients and 307 controls was analysed. All samples corresponded to Spanish white individuals and were recruited at the Hospital Clínico San Carlos (Madrid, Spain) after obtaining written informed consent. Diagnosis was performed by standard clinical, radiological, endoscopic and histological criteria [167]. Controls were mainly anonymous blood donors and laboratory staff. This study was approved by the Ethics Committee of the Hospital.

### Genotyping

DNA was extracted from fresh peripheral blood leukocytes by a "salting out" procedure with a success rate of 97%.

Nine single nucleotide polymorphisms (SNPs) were genotyped by TaqMan technology under the conditions recommended by the manufacturer (Applied Biosystems, Foster City, CA, USA), three in the *CARD15* gene: rs2066844 (R702W), rs2066845 (G908R) and rs2066847 (1007fsinsC); three in *IL23R*: rs1004819, rs7517847 and rs11209026; and three in *TLR9*: rs352162, rs187084 and rs5743836. Case and control samples were included in the same plate for genotyping. Success genotyping rate was over 96% for all samples.

The three SNPs genotyped in *CARD15* are the classically well-established CD-associated mutations. The *IL23R* polymorphisms rs7517847 and rs11209026 were previously studied by our group [168] based on their strong association with CD [126] and rs1004819 was selected because of their described interaction with *TLR9* [166]. The three *TLR9* SNPs were selected to tag most of the variability present in the gene.

### Statistical analysis

Allelic and carrier frequencies between patients and controls were compared using chi-square tests, except for *CARD15*, which was analysed by evaluating the dose effect of carrying any mutant variant using a test of trend. Genetic interactions were evaluated, first, at the patient group analysing the contingency tables in which *TLR9*

genotypic data were confronted to *CARD15* and *IL23R* data by using chi-square tests. When a significant result was obtained, controls were evaluated to check that no interaction was present and the replication data set was used to corroborate the interaction observed. Finally, the specific case-control comparison was performed. Statistical analyses were carried out using the statistical package Epi Info v. 6.02 (World Health Organization, Geneva, Switzerland).

**Results**

Genotypic frequencies of all the SNPs studied conformed to Hardy-Weinberg expectations.

In Table 1, genotypic frequencies for the different SNPs (*CARD15* variants considered together) are shown. The strong susceptibility effect conferred by the *CARD15* mutations is shown at the bottom of the table. Significant differences between CD patients and controls were also observed when considering the three *IL23R* SNPs studied. A protective effect is caused by carriage of the minor allele of rs7517847 and rs11209026 and a susceptibility effect by carriage of the minor allele of rs1004819, concordant with previous literature. Genetic frequencies of *TLR9* polymorphisms do not significantly differ between patients and controls, although a borderline significance is observed when considering the *TLR9* rs352162.

**Table 1**. Genotypic frequencies of the *CARD15*, *IL23R* and *TLR9* SNPs studied in Crohn´s disease (CD) patients and controls

| **Gene** | **CD** | | **Controls** | |
|---|---|---|---|---|
| | n | % | n | % |
| [1]***CARD15*** | | | | |
| 0 | 224 | 63.6 | 497 | 88.3 |
| 1 | 109 | 31.0 | 64 | 11.4 |
| 2 | 19 | 5.40 | 2 | 0.36 |
| ***IL23R*** | | | | |
| [2]rs7517847 | | | | |
| TT | 145 | 42.8 | 194 | 35.5 |
| TG | 159 | 46.9 | 260 | 47.6 |
| GG | 35 | 10.3 | 92 | 16.8 |
| [3]rs11209026 | | | | |
| GG | 320 | 94.4 | 486 | 88.8 |
| GA | 18 | 5.31 | 60 | 11.0 |
| AA | 1 | 0.29 | 1 | 0.18 |
| [4]rs1004819 | | | | |
| CC | 152 | 38.8 | 268 | 49.6 |
| CT | 202 | 51.5 | 218 | 40.4 |
| TT | 38 | 9.69 | 54 | 10.0 |

**Table 1.** Continued

| Gene | CD | | Controls | |
|---|---|---|---|---|
| | n | % | n | % |
| ***TLR9*** | | | | |
| [5]rs352162 | | | | |
| CC | 115 | 27.6 | 167 | 30.5 |
| TC | 192 | 46.2 | 267 | 48.8 |
| TT | 109 | 26.2 | 113 | 20.7 |
| rs187084 | | | | |
| TT | 158 | 38.0 | 183 | 33.8 |
| TC | 192 | 46.2 | 256 | 47.2 |
| CC | 66 | 15.9 | 103 | 19.0 |
| rs5743836 | | | | |
| TT | 295 | 73.4 | 368 | 72.6 |
| TC | 100 | 24.9 | 127 | 25.0 |
| CC | 7 | 1.74 | 12 | 2.37 |

Allelic ORs are shown for the minor allele.

[1] 0, 1 or 2 denotes the number of *CARD15* CD-associated variants (rs2066844=R702W, rs2066845=G908R or rs2066847=1007fsinsC) present. Trend $p<10^{-5}$. OR for 1 mutation=3.78 95% CI 2.63-5.43 and OR for 2 mutations=21.08 95% CI 5.00-187.46.

[2] Allelic: p=0.0037 OR=0.74 95% CI 0.61-0.91; carrier G: p=0.031 OR=0.74 95% CI 0.55-0.98.

[3] Allelic: p=0.0081 OR=0.51 95% CI 0.29-0.87; carrier A: p=0.0051 OR=0.47 95% CI 0.27-0.83.

[4] Allelic: p=0.016 OR=1.27 95% CI 1.04-1.55; carrier T: p=0.0010 OR=1.56 95% CI 1.18-2.04.

[5] Allelic p=0.066 OR=1.18 95% CI 0.98-1.42; TT genotype: 0.043 OR=1.36 95% CI 1.00-1.86.

Subsequent stratified analyses showed two significant *IL23R*/*TLR9* interactions in the CD patient group which were not present in controls (Figure 1): *IL23R* rs1004819 with *TLR9* rs5743836 and *IL23R* rs7517847 with *TLR9* rs352162. When these genetic interactions were evaluated in a second sample set of patients, a low p-value was only observed when considering *IL23R* rs7517847 with *TLR9* rs352162 (p=0.09). In Figure 2, carriers of the minor allele of *IL23R* rs7517847 stratified by the different *TLR9* rs352162 genotypes are shown for that new sample of patients and controls. It seems that carriage of the minor allele of *IL23R* rs7517847 only confers

protection to a subgroup of patients, those who are homozygous for the major allele of *TLR9* rs352162. Therefore, a case-control comparison was performed considering only that *TLR9* subgroup and a highly significant value was obtained: p=0.00032 OR=0.44 (0.27-0.71) (values calculated considering all the CD patients and controls studied).

We investigated the genetic interaction between *IL23R* rs7517847 and *TLR9* rs352162 in UC patients but a significant result was not observed. Of note, individual significant differences in UC emerged only when studying rs11209026, which also shows a protective effect. No interaction was observed between *CARD15* and *TLR9* in CD patients (data not shown).

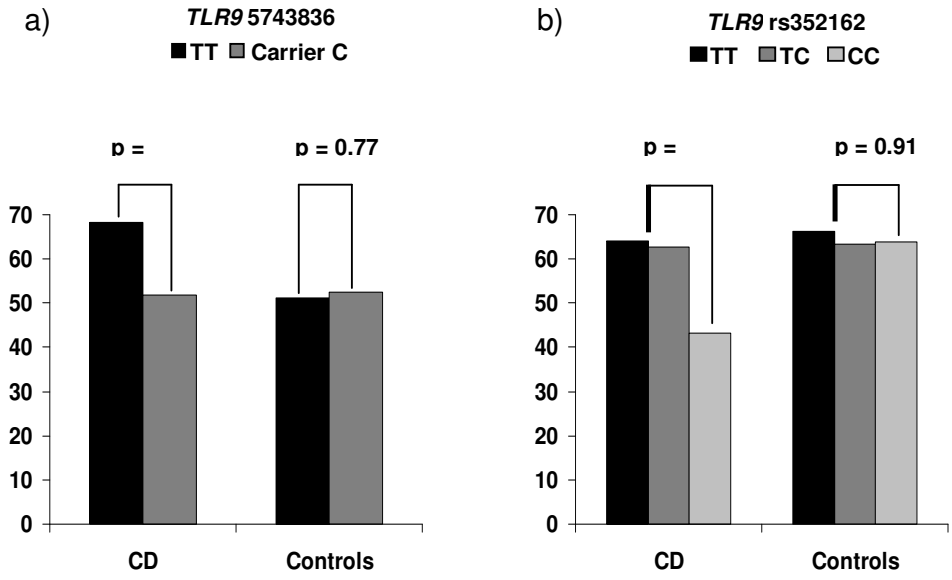



**Figure 1**. Frequency (%) of carriers of the minor allele of the *IL23R* polymorphism a) rs1004819 and b) rs7517847, stratified by the specified *TLR9* genotype.
P values result from the 2*2 *TLR9/IL23R* table. In *TLR9* rs352162, TT+TC are grouped because a similar frequency of *IL23R* is observed in both *TLR9* genotypes

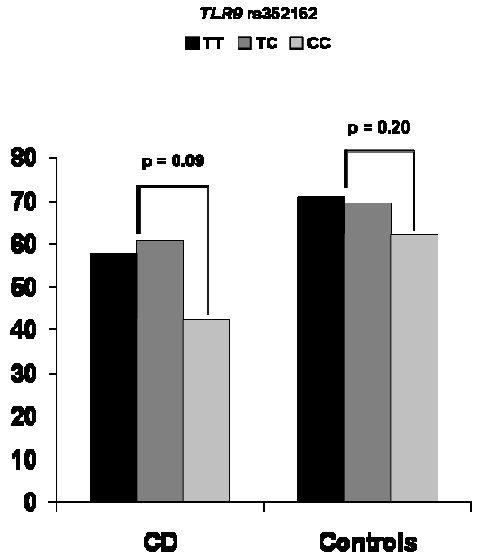



**Figure 2.** Frequency (%) of carriers of the minor allele of the *IL23R* rs7517847 polymorphism stratified by the specified *TLR9* genotype in the replication sample.

**Discussion**

We have investigated the possible interaction between *TLR9* and *IL23R* polymorphisms in order to contribute to CD susceptibility. First, we explored the role of each studied SNP separately to know their individual impact on the disease and, as previously described, the three *IL23R* polymorphisms rs1004819, rs7517847 and rs11209026 show significant association with CD, the minor allele of the first SNP showing a susceptibility effect and the other two showing protection. In *TLR9*, only a borderline significance was observed for rs352162. Next, we performed *IL23R/TLR9* interaction analyses and we observed that the protective effect described for *IL23R* rs7517847 only affected to a specific group of CD patients, those characterized by being homozygous for the major allele (allele C) of *TLR9* rs352162. Obviously, when considering this CD subgroup, a stronger effect than when considering the overall patients was observed: OR=0.44 95% CI 0.27-0.71 *vs*. OR=0.74 95% CI 0.55-0.98, respectively. This low OR explains why a still significant value is observed when considering all CD patients without stratification by *TLR9*.

The genetic interaction observed in our sample contrasts apparently with that described by Török et al. [166], who found the strongest epistasis between *TLR9* rs5743836 and *IL23R* rs1004819. However, it must be noted that in their German patients, *IL23R* rs1004819 showed the strongest association with CD [169], but the top associated SNP in Spanish patients is rs7517847, the one which interacts with *TLR9* in our sample. Although those authors claim that independent effects are observed for the *IL23R* SNPs included in this work, haplotypic analysis in our sample seems to indicate that the susceptibility caused by rs1004819 could be a consequence of its linkage disequilibrium with rs7517847. Therefore, variations in the linkage disequilibrium pattern between Spain and Germany could be underlying the apparent discrepancy in the interacting SNPs. Discrepancies regarding the most associated SNP were also previously evidenced among other populations (see pages 5-6 in [169]).

We also evaluated the interaction between *TLR9* polymorphisms and the three classically CD-associated *CARD15* mutations, but no significant results were obtained. Although our results do not replicate the epistasis described by Törok et al. [166], this was observed neither in New Zealand [165] nor British [170] patients. Therefore, new studies are mandatory to clarify if this is a true interaction or it was a spurious result.

*IL23R* polymorphisms seem to also play a role in UC susceptibility, although causing lower effect than to CD [169]. However, the epistatic effect between *IL23R* and *TLR9* observed in our CD samples does not seem to be modulating UC risk.

As before mentioned, the strong protective effect (low OR) against CD susceptibility caused by the combined presence of those specific genotypes in *IL23R* rs7517847 and *TLR9* rs352162, favours that a protective effect is also observed when considering *IL23R* rs7517847 individually. However, the situation can be different depending on which percentage of individuals with the specific genotype in one SNP also carries the specific genotype to confer risk/protection in the other SNP and depending on the magnitude of the effect. Therefore, it seems reasonable to perform more extensive epistatic studies including SNPs in different genes, although individual signals of association would have not been previously reported. Of course, provided numerous genes are included, the problem lies in the lack of an easy methodology to carry out this kind of analysis as well as in the high sample size needed due to the stringent statistical correction required.

With this work, the presence of an interaction between *TLR9* and *IL23R* polymorphisms contributing to CD susceptibility is confirmed. This raises important questions as the functional relevance of this genetic interaction. As Török et al [166] state, IL-23 production is induced by innate immune system cells through stimulation of pattern recognition receptors by bacterial components. However, the role of IL-23 as amplifier of a Th-17 immune response can also occur in an antigen-independent manner. Therefore, further research is needed to find out the specific nature of the observed epistasis.

Other relevant issue is the clinical implication resulting from this genetic interaction. Specially, it would be interesting to investigate the possible influence of these polymorphisms in relation to the anti-IL12/IL23 therapy, which has been tested in CD [171]. In addition, this kind of study could be extended to other diseases, as psoriasis, in which this biological treatment is effective [172].

**Acknowledgements**

We are most grateful to Carmen Martínez Cuervo and M. Ángel García Martínez for their expert technical assistance. Concepción Núñez has a FIS contract (CP08/0213) and Elena Urcelay works for the “Fundación para la Investigación Biomédica-Hospital Clínico San Carlos”.

**References**

1. Barrett JC, Hansoul S, Nicolae DL, Cho JH, Duerr RH, Rioux JD, et al. Genome-wide association defines more than 30 distinct susceptibility loci for Crohn's disease. Nat Genet 40(8):955-62, 2008.

2. Takahashi M, Kimura A. HLA and CTLA4 polymorphisms may confer a synergistic risk in the susceptibility to Graves' disease. J Hum Genet 55(5):323-6, 2010.

3. Huang CH, Cong L, Xie J, Qiao B, Lo SH, Zheng T. Rheumatoid arthritis-associated gene-gene interaction network for rheumatoid arthritis candidate genes. BMC Proc 3 Suppl 7:S75, 2009.

4. McGovern DP, Rotter JI, Mei L, Haritunians T, Landers C, Derkowski C, et al. Genetic epistasis of IL23/IL17 pathway genes in Crohn's disease. Inflamm Bowel Dis 15(6):883-9, 2009.

5. Petermann I, Huebner C, Browning BL, Gearry RB, Barclay ML, Kennedy M, et al. Interactions among genes influencing bacterial recognition increase IBD risk in a population-based New Zealand cohort. Hum Immunol 70(6):440-6, 2009.

6. Torok HP, Glas J, Endres I, Tonenchi L, Teshome MY, Wetzke M, et al. Epistasis between Toll-like receptor-9 polymorphisms and variants in NOD2 and IL23R modulates susceptibility to Crohn's disease. Am J Gastroenterol 104(7):1723-33, 2009.

7. Lennard-Jones JE. Classification of inflammatory bowel disease. Scand J Gastroenterol Suppl 170:2-6; discussion 16-9, 1989.

8. Marquez A, Mendoza JL, Taxonera C, Diaz-Rubio M, De La Concha EG, Urcelay E, et al. IL23R and IL12B polymorphisms in Spanish IBD patients: no evidence of interaction. Inflamm Bowel Dis 14(9):1192-6, 2008.

9. Duerr RH, Taylor KD, Brant SR, Rioux JD, Silverberg MS, Daly MJ, et al. A genome-wide association study identifies IL23R as an inflammatory bowel disease gene. Science 314(5804):1461-3, 2006.

10. Glas J, Seiderer J, Wetzke M, Konrad A, Torok HP, Schmechel S, et al. rs1004819 is the main disease-associated IL23R variant in German Crohn's disease patients: combined analysis of IL23R, CARD15, and OCTN1/2 variants. PLoS One 2(9):e819, 2007.

11. Cotterill L, Payne D, Levinson S, McLaughlin J, Wesley E, Feeney M, et al. Replication and meta-analysis of 13,000 cases defines the risk for interleukin-23

receptor and autophagy-related 16-like 1 variants in Crohn's disease. Can J Gastroenterol 24(5):297-302, 2010.

12. Sandborn WJ, Feagan BG, Fedorak RN, Scherl E, Fleisher MR, Katz S, et al. A randomized trial of Ustekinumab, a human interleukin-12/23 monoclonal antibody, in patients with moderate-to-severe Crohn's disease. Gastroenterology 135(4):1130-41, 2008.

13. Cingoz O. Ustekinumab. MAbs 1(3):216-21, 2009.

**III.** ***Th17-related genes and celiac disease susceptibility***

Luz María Medrano[1], Manuel García-Magariños[2,3], Bárbara Dema[1], Laura Espino[1], Carlos Maluenda[4], Isabel Polanco[5], M. Ángeles Figueredo[1], Miguel Fernández-Arquero[1] and Concepción Núñez[1]

[1] UGC de Inmunología, Hospital Clínico San Carlos, Instituto de Investigación Sanitaria del Hospital Clínico San Carlos (IdISSC), Madrid, Spain
[2] Unidade de Xenética, Instituto de Medicina Legal and Departamento de Anatomía Patológica y Ciencias Forenses, Facultade de Medicina, Universidade de Santiago de Compostela, Spain
[3] Departamento de Estadística e IO, Universidad Pública de Navarra, Pamplona, Spain
[4] Servicio de Pediatría, Hospital Clínico San Carlos, Madrid, Spain
[5] Servicio de Gastroenterología Pediátrica, Hospital La Paz, Madrid, Spain

**PLoS One. 2012;7(2):e31244.**

## Abstract

Th17 cells are known to be involved in several autoimmune or inflammatory diseases. In celiac disease (CD), recent studies suggest an implication of those cells in disease pathogenesis. We aimed at studying the role of genes relevant for the Th17 immune response in CD susceptibility. A total of 101 single nucleotide polymorphisms (SNPs), mainly selected to cover most of the variability present in 15 Th17-related genes (*IL23R, RORC*, *IL6R, IL17A, IL17F, CCR6, IL6, JAK2, TNFSF15, IL23A, IL22, TBX21, SOCS3, IL12RB1* and *IL17RA*), were genotyped in 735 CD patients and 549 ethnically matched healthy controls. Case-control comparisons for each SNP and for the haplotypes resulting from the SNPs studied in each gene were performed using chi-square tests. Gene-gene interactions were also evaluated following different methodological approaches. No significant results emerged after performing the appropriate statistical corrections. Our results seem to discard a relevant role of Th17 cells on CD risk.


**Key words:** celiac disease susceptibility, gene-gene interactions, Th17 cells

**Introduction**

Celiac disease (CD) is an immune related disease mainly characterized by intestinal inflammation after gluten ingestion in genetically susceptible individuals. CD has been traditionally considered a Th1-mediated disease. However, accumulating evidence about the relevant role of the novel Th17 immune response in several autoimmune diseases[173] opened the possibility towards an involvement of this immunological pathway in CD pathogenesis. These cells seem to be involved in protective responses against extracellular pathogens but they can contribute to chronic inflammation and autoimmunity when dysregulated.

Th17 cells develop from naïve CD161+ CD4+ T cells upon stimulation with particular immunological stimulus, especifically, transforming growth factor beta (TGF-β), interleukin (IL)-23, IL-1β or IL-6 [174]. This induces several transcription factors, mainly the RAR-related orphan receptor C (RORC), which in turn activates IL-17A and IL-17F transcription, the distinctive effector cytokines of this subset of T cells. Production of IL-21, IL-22 and IL-26 also characterizes this specific response, besides the surface markers C-C chemokine receptor type 6 (CCR6) and IL-23 receptor (IL-23R).

Studies based on murine models of several autoimmune diseases, as multiple sclerosis (autoimmune encephalomyelitis, EAE), rheumatoid arthritis (collagen-induced arthritis, CIA) and inflammatory bowel disease (experimental colitis), provided the first evidence about a role of Th17 cells in those conditions [175, 176]. This idea was later supported by case-control studies, which associated genetic variants in *IL23R* with susceptibility to Crohn´s disease, psoriasis and ankylosing spondylitis[126, 177, 178]. Nowadays, the Th17 immune response is considered as a relevant player in several autoimmune or inflammatory diseases. IL-17 mRNA or protein have been detected in biological fluids or the specific affected tissue in several autoimmune disorders [179] and genetic studies associated genes coding important Th17 related products with several diseases [180]. In addition, epistasis between *IL23R* and other Th17 related genes has been reported: with *IL2/IL21* in UC [181] and with *IL17A* and *IL17RA* in Crohn´s disease [164].

In 2008, a putative implication of the Th17 immune response in CD pathogenesis was suggested from two studies following different approaches. Our research group detected a significant association between a genetic polymorphism in the *IL23R* gene and CD[182] and Harris et al. found higher production of IL-23 after

stimulation of human monocytes derived from CD patients with peptic fragments of wheat gliadin[183]. Subsequently, genetic linkage with the *IL23R* region was observed in Finnish families, although this result was not replicated in Hungarian pedigrees and no association with *IL23R* polymorphisms was observed in Finnish, Hungarian or Italian CD samples[184]. In addition, increased expression of several Th17-related cytokines or products was detected in patients with active CD [116, 185] and very recently, gluten-specific IL-17A-producing cells have been found in the duodenum of CD patients [117], which supports a role of Th17 cells in CD pathogenesis.

Despite these results observed in CD, the role of the Th17 cells on this disease is still not well defined. We aimed at shedding more light upon this issue by performing an extensive genetic study including many genes coding distinctive cytokines, markers or transcription factors involved in the Th17 response. We will evaluate the individual influence of those genes on CD susceptibility and also the possible contribution of gene-gene interactions. Previous genome wide association studies (GWAS) did not find association with CD susceptibility of any Th17-related gene [57-59] (with exception of the *IL2/IL21* locus, also involved in other processes), but we consider that a different scenario could emerge with this study: we cover most of the variability present in the studied genetic regions and we will evaluate the genetic interactions between the included polymorphisms, which has been proved as a valid approach to detect new susceptibility variants [166, 186].

## Materials and Methods

### Ethics Statement

This study was approved by the ethical committee (CEIC) of Hospital Clínico San Carlos. Samples were obtained after obtaining written informed consent.

### Subjects

A total of 735 CD patients and 549 ethnically matched healthy controls were included in the initial study. A second sample set consisting of 294 CD patients and 475 controls was used for additional analysis. All these samples correspond to unrelated Spanish white individuals. CD patients were diagnosed following the European Society for Paediatric Gastroenterology and Nutrition (ESPGAN), 97% are positive for HLA-DQ2 and/or HLA-DQ8. Controls correspond mainly to blood donors

and laboratory staff. CD samples were consecutively collected in two centres of the same region (Hospital La Paz and Hospital Clínico San Carlos, Madrid) and controls were collected at the Hospital Clínico San Carlos. A written informed consent was obtained from all the participants in the study, which was approved by the Ethics Committee of the Hospital Clínico San Carlos.

**Markers and genotyping**

We selected genes with a known functional role in the Th17 immune response. Accordingly, fifteen genes were studied: *IL23R, RORC*, *IL6R, IL17A, IL17F, CCR6, IL6, JAK2, TNFSF15, IL23A, IL22, STAT3*, *TBX21, SOCS3, IL12RB1* and *IL17RA*. For all these genes except *IL6R, JAK2* and *STAT3*, single nucleotide polymorphisms (SNPs) were selected following the "aggressive tagging" option present in the Haploview program with genetic data downloaded from the HapMap Project (http://hapmap.ncbi.nlm.nih.gov) (50kb upstream and downstream of the transcription initiation site). To increase statistical power, only markers with a minor allele frequency (MAF)>10% were included. In addition, SNPs located in those genes which code nonsynonymous changes or were previously associated with some autoimmune disease were also analysed independently of their MAF. In *STAT3* and *JAK2*, only two SNPs previously associated with Crohn´s disease, which share some susceptibility factors with CD, were included: rs744166 and rs10758669, respectively; and in *IL6R* we studied one functional polymorphism, rs8192284. SNPs located in *IL6* and *IL6R* and two SNPs in *IL23R*, rs11209026 and rs7517847, were analysed in previous works [182, 187], which included most of the samples initially analysed in this study, but their data were used to evaluate genetic interactions with other Th17 related genes.

A total of 101 SNPs were initially studied (Supplementary Table 1). All of them were genotyped by Veracode technology performed at the National Genotyping Center (http://www.cegen.org), except those that failed (rs10494269, rs9395767, rs608137, rs6927645, rs273506 and rs2241044) and those located in the *IL6*, *IL6R* and *TBX21* genes, which were genotyped with specific TaqMan assays. Two SNPs (rs11209026 and rs7517847, both in the *IL23R* gene) were genotyped by those two technologies and identical results were obtained.

Additional analysis included the study by TaqMan technology of rs12070470, in *IL23R*.

**Statistical analysis**

Deviations from Hardy-Weinberg proportions were assessed in all the SNPs studied.

A case-control analysis using chi-square tests was performed for each individual SNP and for the haplotypes resulting from the SNPs studied in each gene.

Interactions between genes were evaluated following four different approaches: logistic regression, random forests (RF), classification and regression trees (CART) and multifactor dimensionality reduction (MDR).

**Results**

Three SNPs showed deviation from Hardy-Weinberg proportions and were eliminated from the study: rs2064331 (*IL17F*), rs10878804 (*IL22*) and rs9645406 (*RORC*).

The comparison of genotypic frequencies between cases and controls for all the SNPs analysed achieved a nominal significant value in twelve polymorphisms located in eight different genetic regions (Table 1). Although none of them withstand Bonferroni correction, we tried to replicate associations involving *SOCS3* and *IL23R* using a second sample set. These two genes show the lowest case-control p-values in the present analysis and additionally some SNP with a nominal significance in previous CD GWAS [59].

**Table 1.** Genetic polymorphisms which showed a nominal significant value after case-control comparisons (in decreasing significance).

| GENE | SNP | GENOTYPE | p | OR | 95% CI |
|---|---|---|---|---|---|
| *SOCS3* | rs4969170 | AA | 0.0018 | 0.59 | 0.42-0.84 |
| *IL23R* | rs7528924 | GG | 0.0057 | 2.11 | 1.19-3.74 |
| *TNFSF15* | rs17219926 | CC | 0.0103 | 1.43 | 1.08-1.89 |
| *IL6* | rs2069827 | GT+TT | 0.016 | 1.51 | 1.06-2.14 |
| *IL22* | rs11611206 | AA | 0.019 | 0.39 | 0.16-0.93 |
| *IL23R* | rs11209026 | AG+GG | 0.026 | 1.42 | 1.03-1.97 |
| *RORC* | rs1521186 | AA+AG | 0.027 | 1.31 | 1.02-1.67 |
| *IL22* | rs11177131 | CT+TT | 0.034 | 0.76 | 0.58-0.99 |
| *IL17A* | rs8193036 | CT+TT | 0.034 | 0.60 | 0.36-0.99 |
| *IL6* | rs1800795 | CG+CC | 0.037 | 1.26 | 1.01-1.58 |
| *TNFSF15* | rs6478108 | CT+CC | 0.043 | 0.79 | 0.63-1.00 |
| *CCR6* | rs3798315 | TT | 0.044 | 4.17 | 0.90-38.84 |

ORs are referred to the mutant genotype or carrier of the mutant allele (specified below "genotype").

The initial *IL23R* data analysis also evidenced one haplotype significantly associated with CD susceptibility (rs4655683-rs10889667-rs1569922-rs790632-rs7517847-rs10489629-rs7528924-rs2201841-rs4655530-rs11209026-rs6682033-rs6693831, G-C-C-C-T-A-G-T-A-G-A-C): 9.2% in CD patients vs. 6.3% in controls (p=0.0067). For replication purposes, the SNP rs12070470, highly correlated with that haplotype ($r^2$=1 according to http://hapmap.ncbi.nlm.nih.gov/) was studied in the second sample set instead of the 12 SNPs initially considered.

No significant associations involving *IL23R* were observed in the replication set. Regarding the SNP rs4969170, in the *SOCS3* gene, a significant association was observed pooling the original and the replication sets: p=0.0012 OR= 0.64 95% CI 0.49-0.84 (Table 2). Statistical power limitations probably precluded us to obtain a significant result in the replication set.

**Table 2.** Genotypic data (N (%)) for rs4969170 in the original and the replication sets.

| | **Original set[1]** | | **Replication set[2]** | |
|---|---|---|---|---|
| | **CD** (N=732) | **Controls** (N=551) | **CD** (N=294) | **Controls** (N=462) |
| GG | 292 (39.9) | 212 (38.5) | 124 (42.2) | 198 (42.9) |
| AG | 368 (50.3) | 253 (45.9) | 138 (46.9) | 199 (43.1) |
| AA | 72 (9.8) | 86 (15.6) | 32 (10.9) | 65 (14.1) |

AA genotype: 1: p=0.0018 OR=0.59 95% CI 0.42-0.84; 2, p=0.20 95% CI OR=0.75 (0.46- 1.20).

No consensus exists as to the best methodology to evaluate epistasis. Therefore, we used four different statistical methods to evaluate genetic interactions between all the studied polymorphisms located in different genes. We did not find statistically significant results with any methodological approach.

**Discussion**

With the development of genome wide association studies (GWAS), the number of discovered genes involved in CD susceptibility has highly increased. However, the percentage of disease heritability explained has not experienced such an increase. Genetic variants not included in GWAS and genetic interactions could be underlying some missing heritability. We bear this in mind when studying the relevance of the Th17 immune response on CD susceptibility. We performed an extensive case-control study including fifteen genes which code relevant factors involved in that immune response. Tag SNPs were selected to cover most of the variability present in each gene, with exception of *IL6R*, *STAT3* and *JAK2*. SNPs coding nonsynonymous changes or those previously associated with other autoimmune diseases were also included in order to increase the *a priori* probability of obtaining a significant result. Additionally, we evaluated the possibility that interactions between the studied genes were involved in disease susceptibility. Our results seem to discard a relevant role of Th17 cells on CD risk, since no significantly associated SNP or gene-gene interaction was consistently observed, with the only exception of rs4969170, located in *SOCS3*, which deserves further research. However, although *SOCS3* is later confirmed, its functional role must be elucidated, since it is involved in different functional pathways and it would be expected that more than one Th17 gene was associated with CD susceptibility, as it has been observed with other Th17-mediated diseases.

The discovery of the IL-23 cytokine prompted the re-examination of the dominant Th response in many autoimmune diseases, primarily in those considered as skewed towards a Th1 phenotype. Studies based on murine models of multiple sclerosis, rheumatoid arthritis and inflammatory bowel disease related these conditions with a Th17 response. However, a similar conclusion is not drawn from GWAS results [53, 188, 189]. Although several Th17-related genes have been associated with IBD and RA, the list of MS susceptibility genes does not suggest a Th17 related etiology. This intriguing issue is probably far away from being answered. Nowadays it seems clear that Th immune responses are not independent and plasticity exists between Th cell subsets. A shift between Th1 and Th17 can occur during the inflammatory process and it is possible to speculate that the relative contribution and the timing of each subset will determine which genes would be involved in disease risk. Moreover, the cytokine microenvironment can determine the shift towards a specific immune response. From this point of view, genetics could not be so relevant if other

compensatory mechanisms exist. This evidence, as previously suggested, that overlap between autoimmune diseases must be observed with caution. Th17 cells seem to mediate several autoimmune diseases but their impact in disease etiology seems to be different.

In summary, gene expression studies link CD pathogenesis to Th17 cells, but we evidenced that polymorphisms in Th17-related genes do not seem to be crucial for disease development. This is concordant with observations on MS. Although, in general, genetic data provide clues that ratified by functional studies unravel disease pathogenesis, this time it makes necessary to do somehow the other way around, with the special difficulty of explaining the divergent genetic results observed in different immune mediated diseases. Therefore, much more work is expected in this field.

**Acknowledgements**

We are most grateful to Carmen Martínez Cuervo and M. Ángel García Martínez for her expert technical assistance.

**Supplementary Table 1**. Genes and SNPs studied ordered by chromosome and position.

| GENE | SNP | CHROMOSOME | POSITION |
|---|---|---|---|
| *IL23R* | rs4655683 | 1 | 67384201 |
| *IL23R* | rs10889667 | 1 | 67427795 |
| *IL23R* | rs1569922 | 1 | 67437551 |
| *IL23R* | rs790632 | 1 | 67451510 |
| *IL23R* | rs7517847 | 1 | 67454257 |
| *IL23R* | rs10489629 | 1 | 67460937 |
| *IL23R* | rs7528924 | 1 | 67461624 |
| *IL23R* | rs2201841 | 1 | 67466790 |
| *IL23R* | rs4655530 | 1 | 67476319 |
| *IL23R* | rs11209026 | 1 | 67478546 |
| *IL23R* | rs6682033 | 1 | 67481258 |
| *IL23R* | rs6693831 | 1 | 67493455 |
| *RORC* | rs3828057 | 1 | 150046801 |
| *RORC* | rs12045886 | 1 | 150048714 |
| *RORC* | rs10494269 | 1 | 150048958 |
| *RORC* | rs939595 | 1 | 150050312 |
| *RORC* | rs1521186 | 1 | 150051171 |
| *RORC* | rs6693413 | 1 | 150058934 |
| *RORC* | rs949969 | 1 | 150059438 |
| *RORC* | rs4845604 | 1 | 150068304 |
| *RORC* | rs17582155 | 1 | 150070837 |
| *RORC* | rs6587622 | 1 | 150081029 |
| *RORC* | rs9645406 | 1 | 150083696 |
| *RORC* | rs16833584 | 1 | 150083858 |
| *IL6R* | rs8192284 | 1 | 152693594 |
| *IL17A* | rs13208597 | 6 | 52152468 |
| *IL17A* | rs9395767 | 6 | 52154729 |
| *IL17A* | rs4711998 | 6 | 52158312 |
| *IL17A* | rs8193036 | 6 | 52158452 |
| *IL17A* | rs2275913 | 6 | 52158992 |
| *IL17A* | rs7747909 | 6 | 52162208 |

| | | | |
|---|---|---|---|
| *IL17A* | rs1974226 | 6 | 52163294 |
| *IL17A/IL17F* | rs4715287 | 6 | 52174606 |
| *IL17A/IL17F* | rs12528203 | 6 | 52175294 |
| *IL17F* | rs2064331 | 6 | 52206327 |
| *IL17F* | rs13209590 | 6 | 52207610 |
| *IL17F* | rs641701 | 6 | 52209124 |
| *IL17F* | rs11465553 | 6 | 52209717 |
| *IL17F* | rs7771466 | 6 | 52212068 |
| *IL17F* | rs12201582 | 6 | 52212648 |
| *IL17F* | rs608137 | 6 | 52220675 |
| *CCR6* | rs2301436 | 6 | 167357978 |
| *CCR6* | rs150110 | 6 | 167431756 |
| *CCR6* | rs9364891 | 6 | 167436309 |
| *CCR6* | rs6927645 | 6 | 167450292 |
| *CCR6* | rs3798315 | 6 | 167464699 |
| *CCR6* | rs3093012 | 6 | 167468271 |
| *CCR6* | rs3093010 | 6 | 167468597 |
| *CCR6* | rs3093007 | 6 | 167469765 |
| *CCR6* | rs17860852 | 6 | 167470814 |
| *CCR6* | rs4710187 | 6 | 167475082 |
| *CCR6* | rs4710189 | 6 | 167476314 |
| *CCR6* | rs11575089 | 6 | 167477189 |
| *IL6* | rs2069827 | 7 | 22731981 |
| *IL6* | rs1800795 | 7 | 22733170 |
| *IL6* | rs2069840 | 7 | 22735097 |
| *JAK2* | rs10758669 | 9 | 4971602 |
| *TNFSF15* | rs6478106 | 9 | 116585487 |
| *TNFSF15* | rs6478108 | 9 | 116598524 |
| *TNFSF15* | rs7865494 | 9 | 116616300 |
| *TNFSF15* | rs17219926 | 9 | 116619674 |
| *TNFSF15* | rs1322063 | 9 | 116625303 |
| *IL23A* | rs11613055 | 12 | 55011186 |
| *IL23A* | rs11171806 | 12 | 55019798 |

| | | | |
|---|---|---|---|
| *IL22* | rs741346 | 12 | 66917607 |
| *IL22* | rs10878796 | 12 | 66921463 |
| *IL22* | rs11177131 | 12 | 66923715 |
| *IL22* | rs1182844 | 12 | 66927799 |
| *IL22* | rs976748 | 12 | 66929861 |
| *IL22* | rs12100005 | 12 | 66940909 |
| *IL22* | rs11611206 | 12 | 66954713 |
| *IL22* | rs10878804 | 12 | 66955524 |
| *IL22* | rs1383119 | 12 | 66958414 |
| | rs13377617 | 12 | 67004760 |
| *STAT3* | rs744166 | 17 | 37767727 |
| *TBX21* | rs10514934 | 17 | 43167123 |
| *TBX21* | rs11657388 | 17 | 43175706 |
| *TBX21* | rs4794067 | 17 | 43163827 |
| *SOCS3* | rs4789588 | 17 | 73844676 |
| *SOCS3* | rs6501199 | 17 | 73846381 |
| *SOCS3* | rs11077357 | 17 | 73855196 |
| *SOCS3* | rs8069976 | 17 | 73861445 |
| *SOCS3* | rs4969170 | 17 | 73872133 |
| *SOCS3* | rs9900564 | 17 | 73889053 |
| *SOCS3* | rs4366775 | 17 | 73893674 |
| *IL12RB1* | rs436857 | 19 | 18058635 |
| *IL12RB1* | rs1870063 | 19 | 18031962 |
| *IL12RB1* | rs3761041 | 19 | 18044917 |
| *IL12RB1* | rs429774 | 19 | 18047752 |
| *IL12RB1* | rs374326 | 19 | 18062266 |
| *IL12RB1* | rs273506 | 19 | 18082647 |
| *IL17RA* | rs5994155 | 22 | 15929526 |
| *IL17RA* | rs9606603 | 22 | 15941513 |
| *IL17RA* | rs5748864 | 22 | 15952941 |
| *IL17RA* | rs13053889 | 22 | 15960394 |
| *IL17RA* | rs9606615 | 22 | 15960813 |
| *IL17RA* | rs2241044 | 22 | 15961838 |

| | | | |
|---|---|---|---|
| *IL17RA* | rs721930 | 22 | 15965808 |
| *IL17RA* | rs2241049 | 22 | 15967680 |
| *IL17RA* | rs879575 | 22 | 15969567 |
| *IL17RA* | rs2895332 | 22 | 15971089 |

**References**

1. Oukka M (2008) Th17 cells in immunity and autoimmunity. Ann Rheum Dis 67 Suppl 3: iii26-29.

2. Volpe E, Servant N, Zollinger R, Bogiatzi SI, Hupe P, et al. (2008) A critical function for transforming growth factor-beta, interleukin 23 and proinflammatory cytokines in driving and modulating human T(H)-17 responses. Nat Immunol 9: 650-657.

3. Cua DJ, Sherlock J, Chen Y, Murphy CA, Joyce B, et al. (2003) Interleukin-23 rather than interleukin-12 is the critical cytokine for autoimmune inflammation of the brain. Nature 421: 744-748.

4. Yen D, Cheung J, Scheerens H, Poulet F, McClanahan T, et al. (2006) IL-23 is essential for T cell-mediated colitis and promotes inflammation via IL-17 and IL-6. J Clin Invest 116: 1310-1316.

5. Duerr RH, Taylor KD, Brant SR, Rioux JD, Silverberg MS, et al. (2006) A genome-wide association study identifies IL23R as an inflammatory bowel disease gene. Science 314: 1461-1463.

6. Burton PR, Clayton DG, Cardon LR, Craddock N, Deloukas P, et al. (2007) Association scan of 14,500 nonsynonymous SNPs in four diseases identifies autoimmunity variants. Nat Genet 39: 1329-1337.

7. Cargill M, Schrodi SJ, Chang M, Garcia VE, Brandon R, et al. (2007) A large-scale genetic association study confirms IL12B and leads to the identification of IL23R as psoriasis-risk genes. Am J Hum Genet 80: 273-290.

8. Steinman L (2007) A brief history of T(H)17, the first major revision in the T(H)1/T(H)2 hypothesis of T cell-mediated tissue damage. Nat Med 13: 139-145.

9. Lees CW, Barrett JC, Parkes M, Satsangi J (2011) New IBD genetics: common pathways with other diseases. Gut Epub ahead of print.

10. Glas J, Stallhofer J, Ripke S, Wetzke M, Pfennig S, et al. (2009) Novel genetic risk markers for ulcerative colitis in the IL2/IL21 region are in epistasis with IL23R and suggest a common genetic background for ulcerative colitis and celiac disease. Am J Gastroenterol 104: 1737-1744.

11. McGovern DP, Rotter JI, Mei L, Haritunians T, Landers C, et al. (2009) Genetic epistasis of IL23/IL17 pathway genes in Crohn's disease. Inflamm Bowel Dis 15: 883-889.

12. Nunez C, Dema B, Cenit MC, Polanco I, Maluenda C, et al. (2008) IL23R: a susceptibility locus for celiac disease and multiple sclerosis? Genes Immun 9: 289-293.
13. Harris KM, Fasano A, Mann DL (2008) Cutting edge: IL-1 controls the IL-23 response induced by gliadin, the etiologic agent in celiac disease. J Immunol 181: 4457-4460.
14. Einarsdottir E, Koskinen LL, Dukes E, Kainu K, Suomela S, et al. (2009) IL23R in the Swedish, Finnish, Hungarian and Italian populations: association with IBD and psoriasis, and linkage to celiac disease. BMC Med Genet 10: 8.
15. Monteleone I, Sarra M, Del Vecchio Blanco G, Paoluzi OA, Franze E, et al. (2011) Characterization of IL-17A-producing cells in celiac disease mucosa. J Immunol 184: 2211-2218.
16. Castellanos-Rubio A, Santin I, Irastorza I, Castano L, Carlos Vitoria J, et al. (2009) TH17 (and TH1) signatures of intestinal biopsies of CD patients in response to gliadin. Autoimmunity 42: 69-73.
17. Fernandez S, Molina IJ, Romero P, Gonzalez R, Pena J, et al. (2011) Characterization of gliadin-specific Th17 cells from the mucosa of celiac disease patients. Am J Gastroenterol 106: 528-538.
18. van Heel DA, Franke L, Hunt KA, Gwilliam R, Zhernakova A, et al. (2007) A genome-wide association study for celiac disease identifies risk variants in the region harboring IL2 and IL21. Nat Genet 39: 827-829.
19. Hunt KA, Zhernakova A, Turner G, Heap GA, Franke L, et al. (2008) Newly identified genetic risk variants for celiac disease related to the immune response. Nat Genet 40: 395-402.
20. Dubois PC, Trynka G, Franke L, Hunt KA, Romanos J, et al. (2010) Multiple common variants for celiac disease influencing immune gene expression. Nat Genet 42: 295-302.
21. Bush WS, McCauley JL, DeJager PL, Dudek SM, Hafler DA, et al. (2011) A knowledge-driven interaction analysis reveals potential neurodegenerative mechanism of multiple sclerosis susceptibility. Genes Immun 12: 335-340.
22. Torok HP, Glas J, Endres I, Tonenchi L, Teshome MY, et al. (2009) Epistasis between Toll-like receptor-9 polymorphisms and variants in NOD2 and IL23R modulates susceptibility to Crohn's disease. Am J Gastroenterol 104: 1723-1733.

23. Dema B, Martinez A, Fernandez-Arquero M, Maluenda C, Polanco I, et al. (2009) The IL6-174G/C polymorphism is associated with celiac disease susceptibility in girls. Hum Immunol 70: 191-194.
24. Franke A, McGovern DP, Barrett JC, Wang K, Radford-Smith GL, et al. (2010) Genome-wide meta-analysis increases to 71 the number of confirmed Crohn's disease susceptibility loci. Nat Genet 42: 1118-1125.
25. Stahl EA, Raychaudhuri S, Remmers EF, Xie G, Eyre S, et al. (2010) Genome-wide association study meta-analysis identifies seven new rheumatoid arthritis risk loci. Nat Genet 42: 508-514.
26. Sawcer S, Hellenthal G, Pirinen M, Spencer CC, Patsopoulos NA, et al. (2011) Genetic risk and a primary role for cell-mediated immune mechanisms in multiple sclerosis. Nature 476: 214-219.

2. ESTUDIO DE EXPRESIÓN A NIVEL INTESTINAL

## I. *Th17 related gene expression in CeD and UC patients.*

Luz María Medrano[1], Virginia Pascual[1], Andrés Bodas[2], Juan Luis Mendoza[3], Miguel Fernández-Arquero[1], Natalia López-Palacios[3], Concepción Núñez[1]

[1] UGC de Inmunología, Hospital Clínico San Carlos, Instituto de Investigación Sanitaria del Hospital Clínico San Carlos (IdISSC), Madrid, Spain
[2] Servicio de Pediatría, Hospital Clínico San Carlos, Instituto de Investigación Sanitaria del Hospital Clínico San Carlos (IdISSC), Madrid, Spain
[3] Servicio de Digestivo, Hospital Clínico San Carlos, Instituto de Investigación Sanitaria del Hospital Clínico San Carlos (IdISSC), Madrid, Spain

**(enviado)**

**Abstract**

Introduction: Celiac disease (CeD) and inflammatory bowel disease (Crohn´s disease and ulcerative colitis) are multifactorial diseases characterized by chronic intestinal inflammation resulting from a dysregulated immune response. Their immunopathology is not completely understood yet but last genetic studies showed the involvement of the Th17 cells in IBD and other immune-related diseases. Our objective was to study the involvement of the Th17 pathway in CeD and UC through the quantification of the expression levels of twelve Th17 related genes in intestinal samples from those patients.

Methods: A total of 46 intestinal biopsies were included in our study: 24 duodenal samples from CeD patients, 16 samples from colon or rectum from UC patients and 6 duodenal samples from healthy controls. The expression levels of twelve Th17 related genes (CCR6, GLB1, IL17F, IL2, IL21, IL22, IL23A, IL23R, IL6, JAK2, RHBDD3, SMAD3, STAT3 and TYK2) were determined by quantitative PCR (qPCR).

Results: A higher expression for IL23A, JAK2, IL6 and RHBDD3 was observed in CeD patients on a gluten containing diet compared with controls. The different expression observed for IL6 and RHBDD3 was only present in CeD patients with adult onset. In UC, three genes, CCR6, IL6 and SMAD3, showed higher expression in inflamed samples when compared to samples from healthy tissue from the same patients.

Conclusions: Our results suggest a role of the Th17 immune response in CeD and UC, although different genes seem to be mainly involved in those two diseases. In addition, some differences in the expression profile depending on the age at diagnosis of CeD are also observed, which could arise as a consequence of an age at onset dependent mechanism involved in the pathogenesis of the disease

Keywords: celiac disease, gene expression, intestinal biopsies, Th17 pathway, ulcerative colitis.

**Introduction**

Celiac disease (CeD) is a chronic intestinal disorder caused by ingestion of wheat gluten or similar proteins present in other cereals (mainly rye and barley) in genetic predisposed individuals. Inflammatory bowel disease is also a chronic inflammatory intestinal disease, with two major clinical forms: Crohn´s disease (CD) and ulcerative colitis (UC); it seems to be a consequence of an abnormal immune response to the normal luminal flora in genetic predisposed individuals.[1, 2]

These diseases, as commonly observed in immune-mediated disorders, show a dysregulation of the different subsets of T helper (Th) cells. Depending on the disease, a specific subset of these cells seems to mainly contribute to the pathogenic process. For many years two Th cell subsets were thought to be the most relevant in these diseases: Th1 in CD and CeD [3-5] and Th2 in UC.[4, 6] In the last decades, several genes (*IL23R*, *IL12B*, *STAT3*, *CCR6*, *JAK2*) involved in the Th17 pathway have been associated with susceptibility to IBD and other autoimmune or inflammatory disorders by genome wide association studies (GWAs). [7-10] In addition, gene expression studies also showed an implication of the Th17 cells in the development of IBD.[11-13] The role of the Th17 response in CeD is not as clear. No genes related to the Th17 pathway have been associated with CeD susceptibility in GWAs, with the exception of the *IL2/IL21* locus, but those genes are also involved in other several functions. We previously performed an extensive case-control study including 16 Th17-related genes and we did not observe any significant association with CD susceptibility.[14] However, an altered expression of some Th17 genes in CeD was observed in several works.[15-18]

Th17 cells are characterized by the production of the effector cytokines IL-17A, IL-17F and IL-22. Their differentiation depends on the activation of the JAK2 kinase and the STAT3 transcription factor and these cells are regulated by a combination of cytokines, including IL-6, IL-1B, TGF-β1, IL-23A and the autocrine activity of IL-21. *CCR6* is a surface marker of Th17 cells and contributes to the recruitment of several chemokines. Furthermore, IL-23 regulates the maintenance and expansion of the Th17 cells. [19, 20] The expression of most of the genes coding for these factors has been studied in CD patients,[12, 21, 22] but less information exists in UC patients and

specially in CeD patients. Our aim was to analyze the intestinal expression of numerous Th17 related genes in CeD and UC patients.

## Material and Methods

### Subjects

CeD patients were diagnosed according to the ESPHGAN (European Society for Pediatric Gastroenterology, Hepatology and Nutrition) criteria.[23, 24] Diagnosis of IBD was based on standard clinical, radiologic, endoscopic and histological criteria. [25] A total of 24 duodenal biopsies from CeD patients (20 on a gluten-containing diet and 4 on gluten free diet, GFD) and 16 biopsies from colon or rectum from UC patients (8 paired samples from healthy and inflamed tissue) were included in our study. Six duodenal samples from subjects with normal histology and without immune related diseases were used as controls (4 adults and 2 children). Mucosal inflammation in UC was defined as the presence of macroscopically and microscopically histological signs of disease activity. Two subgroups were defined in CeD depending on the time of diagnosis: pediatric CeD (2 biopsies from patients on GFD and 14 on gluten containing diet) and adult CeD (2 biopsies from patients on GFD and 6 on gluten containing diet). Written informed consent was obtained from all patients and the protocol was approved by the Ethic Committee of Hospital Clínico San Carlos (Madrid, Spain).

### RNA extraction and cDNA synthesis and analysis

All samples were collected in RNAlater (TM RNA Stabilization Reagent, Qiagen) and stored 24-48 hours at 4ºC and thereafter at -70ºC until use. RNA was obtained using the RNeasy Mini Kit (Qiagen, Westburg, Leuden, The Netherlands). Quantity and purity of RNA was determined by spectrophotometry (Nanodrop, Thermo Scientific NanoDrop Products, Wilmington, DE, USA). cDNA was obtained by RT-PCR (High Capacity RNA-to-cDNA Master Mix, Applied Biosystems).

Expression of *CCR6*, *GLB1*, *IL17F*, *IL2*, *IL21*, *IL22*, *IL23A*, *IL23R*, *IL6*, *JAK2*, *RHBDD3*, *SMAD3*, *STAT3* and *TYK2* was measured by quantitative PCR (qPCR) using Custom TaqMan Array Cards (format 384-well microfluidic card, Applied Biosystems). Missing threshold cycles (Ct) values were set as 36.

Gene expression levels were determined by the comparative Ct method. The expression values were expressed according to delta-Ct ($\Delta Ct = \Delta Ct_{target} - \Delta Ct_{endogenous}$), lower ΔCt represents more copy numbers of amplified mRNA. [26] To select the best reference genes, 48 genes including *18s, RPLP0* and *HPRT1* were evaluated using two algorithms: geNorm version 3.5 [27] and NormFinder version 0.953 [28], in the Genex software.

**Statistical Analysis**

Differences between groups of independent samples (pediatric versus adult CeD or CeD patients versus controls) were evaluated by the U-Mann Whitney test and differences between paired samples (healthy versus inflamed tissue) were analyzed by the Wilcoxon test. Depending on the existence of previous information, one or two tailed p-values were considered significant. The SPPS 15.1 program (SPSS, Chicago, USA) was used for statistical analyses.

**Results**

*GLB1* and *TYK2* were used as reference genes because of lower variation in all the groups of samples using geNorm and NormFinder.

We did not observe any significant difference between duodenal samples from adult controls and pediatric controls, therefore we grouped all these samples in order to increase statistical power. First of all, we compared the expression levels for all the studied genes between pediatric and adult CeD patients (Figure 1). *IL6* (p= 0.00036) and *RHBDD3* (p=0.0064) expression was significantly increased in adult CeD patients compared to pediatric patients. *IL6* expression in adult patients was also significantly different from controls (p=0.0087).

Next, we excluded *IL6* and *RHBDD3* and compared all CeD patients (pediatric and adults) with controls (Figure 2). Significant changes in gene expression were observed for *IL23A* (p=0.039) and *JAK2* (p=0.046). No differences were observed between patients on a GFD and on a gluten containing diet; with the exception of *STAT3*, with mRNA levels significantly increased in biopsies from patients on a GFD (p=0.018) (Figure 3).

We also compared paired samples belonging to healthy and inflamed tissue from the same UC patients (Figure 4). We observed increased levels of *CCR6* (p=0.012), *IL6* (p=0.020) and *SMAD3* (p=0.016) in inflamed tissue. Additionally, notorious although non-significant differences were observed for *JAK2* (p=0.055), *IL23R* (p=0.078) and *IL21* (p=0.109), all increased in inflamed samples.

**Discussion**

In the present study the expression of several genes involved in the Th17 pathway was evaluated in CeD and CU patients. We first analyzed the expression in CeD patients stratified according to the time of diagnosis: childhood or adulthood. We observed a significantly lower expression of *IL6* and *RHBDD3* in children compared to adults. It is noteworthy that *Rhbdd3*$^{-}$/*Rhbdd3*$^{-}$ knockout mice have been associated with low levels of IL-6. [29] To our knowledge, this is the first time that a different profile of cytokines has been described for CeD patients depending on the age at onset. The commonly observed differences in the clinical presentation between children and adults could be partially due to differences in the inflammatory response.

When considering all samples from CeD patients on gluten containing diet and compared them with controls, significant differences were observed for *IL23A* and *JAK2* expression. No previous studies related to *JAK2* have been published, but *IL23A* expression was described as increased in patients at diagnosis of CeD and therefore on gluten containing diet compared to patients on a GFD, although that difference did not reach statistical significance. [15]

Evaluation of the expression of these Th17 related genes in CeD patients on a GFD showed an increase of *STAT3* intestinal expression in these patients compared to those on gluten containing diet. No differences with controls were observed for any of these two groups of patients. The low number of samples from CeD patients on GFD compromises the statistical power of our study, therefore it is necessary to increase the sample size in order to validate our results. *STAT3* appears to be an essential signaling molecule for the differentiation of Th17 cells, since the induction of a key cytokine for differentiation of these cells, *IL21*, is dependent on *STAT3* [19]. Moreover, deletion of *STAT3* prevents development of Th17 cells.[30]

Finally, we compared the expression levels of the selected Th17 related genes between samples from inflamed and healthy tissue belonging to the same UC patients. Increased levels of *IL6*, *CCR6* and *SMAD3* in inflamed tissue were observed. IL-6 synergizes with TGF-β in promoting Th17 differentiation with the activation of specific transcription factors which include *SMAD3*.[19, 31-33] *CCR6* is expressed on the surface of Th17 cells and acts as receptor for *CCL20*, a chemokine which contributes to the recruitment of Th17 cells. Moreover, we observed higher levels of expression in inflamed samples for *IL21*, *IL23R* and *JAK2*, although these differences did not reach statistical significance. Previous gene expression studies suggested an involvement of the Th17 immune response in UC. Olsen *et al*. correlated the expression levels of *IL23A* and *IL6* with the degree of inflammation in UC as evaluated by the relative reduction in endoscopic score and UCDAI index.[13] Most recently, a meta-analysis observed significant differences in *IL6* and *CCR6* expression levels between inflamed and un-inflamed colonic biopsies from UC patients [11]. Our data confirm that *IL6* and *CCR6* play a role in mediating inflammation in CU. Further studies are necessary to validate our results in *SMAD3*, but it would be also interesting to further investigate *IL21*, *IL23R* and *JAK2*.

Taking into account that gene expression levels can be different as a consequence of genetic variation, future studies involving the correlation between expression levels in *IL6*, *RHBDD3*, *IL23A* and *JAK2* and the presence of specific variants in those or related genes are necessary to associate these genes with disease development. The altered levels of *IL6* and *RHBDD3* observed only in adult patients would explain the lack of association of these genes in previous GWAs, which mainly include pediatric patients.

Independently of the results of future correlation studies, we observe in this work that a common involvement of Th17 cells in the pathogenesis of both diseases, CeD and UC, exists but important differences seem also to be present because the subset of genes which show an altered expression differs between those two diseases, with the exception of *IL6* and perhaps *JAK2*. Th17 cells seem to be related to a response against extracellular pathogens. In CeD, a dominant role of gluten exists in disease development, therefore the role of pathogens in the origin of the disease is probably weaker than in UC and therefore stronger influence of this kind of response

is observed in UC patients. Further studies are necessary to best define the Th17 differential implication between UC and CeD.

**References**

1.Kaser A, Zeissig S, Blumberg RS. Inflammatory bowel disease. Annu Rev Immunol 28:573-621, 2010.

2.MacDonald TT, Monteleone I, Fantini MC, Monteleone G. Regulation of homeostasis and inflammation in the intestine. Gastroenterology 140(6):1768-75, 2011.

3.Bamias G, Sugawara K, Pagnini C, Cominelli F. The Th1 immune pathway as a therapeutic target in Crohn's disease. Curr Opin Investig Drugs 4(11):1279-86, 2003.

4.Bouma G, Strober W. The immunological and genetic basis of inflammatory bowel disease. Nat Rev Immunol 3(7):521-33, 2003.

5.Ben-Horin S, Green PH, Bank I, Chess L, Goldstein I. Characterizing the circulating, gliadin-specific CD4+ memory T cells in patients with celiac disease: linkage between memory function, gut homing and Th1 polarization. J Leukoc Biol 79(4):676-85, 2006.

6.Targan SR, Karp LC. Defects in mucosal immunity leading to ulcerative colitis. Immunol Rev 206:296-305, 2005.

7.Duerr RH, Taylor KD, Brant SR, Rioux JD, Silverberg MS, Daly MJ, et al. A genome-wide association study identifies IL23R as an inflammatory bowel disease gene. Science 314(5804):1461-3, 2006.

8.Jostins L, Ripke S, Weersma RK, Duerr RH, McGovern DP, Hui KY, et al. Host-microbe interactions have shaped the genetic architecture of inflammatory bowel disease. Nature 491(7422):119-24, 2012.

9.Eyre S, Bowes J, Diogo D, Lee A, Barton A, Martin P, et al. High-density genetic mapping identifies new susceptibility loci for rheumatoid arthritis. Nat Genet 44(12):1336-40, 2012.

10.Tsoi LC, Spain SL, Knight J, Ellinghaus E, Stuart PE, Capon F, et al. Identification of 15 new psoriasis susceptibility loci highlights the role of innate immunity. Nat Genet 44(12):1341-8, 2012.

11.Granlund A, Flatberg A, Ostvik AE, Drozdov I, Gustafsson BI, Kidd M, et al. Whole genome gene expression meta-analysis of inflammatory bowel disease colon

mucosa demonstrates lack of major differences between Crohn's disease and ulcerative colitis. PLoS One 8(2):e56818, 2013.

12.Schmechel S, Konrad A, Diegelmann J, Glas J, Wetzke M, Paschos E, et al. Linking genetic susceptibility to Crohn's disease with Th17 cell function: IL-22 serum levels are increased in Crohn's disease and correlate with disease activity and IL23R genotype status. Inflamm Bowel Dis 14(2):204-12, 2008.

13.Olsen T, Rismo R, Cui G, Goll R, Christiansen I, Florholmen J. TH1 and TH17 interactions in untreated inflamed mucosa of inflammatory bowel disease, and their potential to mediate the inflammation. Cytokine 56(3):633-40, 2011.

14.Medrano LM, Garcia-Magarinos M, Dema B, Espino L, Maluenda C, Polanco I, et al. Th17-related genes and celiac disease susceptibility. PLoS One 7(2):e31244, 2012.

15.Castellanos-Rubio A, Santin I, Irastorza I, Castano L, Carlos Vitoria J, Ramon Bilbao J. TH17 (and TH1) signatures of intestinal biopsies of CD patients in response to gliadin. Autoimmunity 42(1):69-73, 2009.

16.Fernandez S, Molina IJ, Romero P, Gonzalez R, Pena J, Sanchez F, et al. Characterization of gliadin-specific Th17 cells from the mucosa of celiac disease patients. Am J Gastroenterol 106(3):528-38, 2011.

17.Bragde H, Jansson U, Jarlsfelt I, Soderman J. Gene expression profiling of duodenal biopsies discriminates celiac disease mucosa from normal mucosa. Pediatr Res 69(6):530-7, 2011.

18.Plaza-Izurieta L, Castellanos-Rubio A, Irastorza I, Fernandez-Jimenez N, Gutierrez G, Bilbao JR. Revisiting genome wide association studies (GWAS) in coeliac disease: replication study in Spanish population and expression analysis of candidate genes. J Med Genet 48(7):493-6, 2011.

19.Bettelli E, Korn T, Oukka M, Kuchroo VK. Induction and effector functions of T(H)17 cells. Nature 453(7198):1051-7, 2008.

20.Brand S. Crohn's disease: Th1, Th17 or both? The change of a paradigm: new immunological and genetic insights implicate Th17 cells in the pathogenesis of Crohn's disease. Gut 58(8):1152-67, 2009.

21.Bogaert S, Laukens D, Peeters H, Melis L, Olievier K, Boon N, et al. Differential mucosal expression of Th17-related genes between the inflamed colon and ileum of patients with inflammatory bowel disease. BMC Immunol 11:61, 2010.

22.Noble CL, Abbas AR, Lees CW, Cornelius J, Toy K, Modrusan Z, et al. Characterization of intestinal gene expression profiles in Crohn's disease by genome-wide microarray analysis. Inflamm Bowel Dis 16(10):1717-28, 2010.
23.ESPHGAN. Revised criteria for diagnosis of coeliac disease. Report of Working Group of European Society of Paediatric Gastroenterology and Nutrition. Arch Dis Child 65(8):909-11, 1990.
24.Husby S, Koletzko S, Korponay-Szabo IR, Mearin ML, Phillips A, Shamir R, et al. European Society for Pediatric Gastroenterology, Hepatology, and Nutrition guidelines for the diagnosis of coeliac disease. J Pediatr Gastroenterol Nutr 54(1):136-60, 2012.
25.Lennard-Jones JE. Classification of inflammatory bowel disease. Scand J Gastroenterol Suppl 170:2-6; discussion 16-9, 1989.
26.Schmittgen TD, Livak KJ. Analyzing real-time PCR data by the comparative C(T) method. Nat Protoc 3(6):1101-8, 2008.
27.geNorm, version 3.5; Center for Medical Genetics. Ghent University; Ghent, Belgium. 2007.
28.NormFinder, version 0.953. Aarhus University Hospital; Aarhus, Denmark. 2005.
29.Liu J, Liu S, Xia M, Xu S, Wang C, Bao Y, et al. Rhomboid domain-containing protein 3 is a negative regulator of TLR3-triggered natural killer cell activation. Proc Natl Acad Sci U S A 110(19):7814-9, 2013.
30.Yang XO, Panopoulos AD, Nurieva R, Chang SH, Wang D, Watowich SS, et al. STAT3 regulates cytokine-mediated generation of inflammatory helper T cells. J Biol Chem 282(13):9358-63, 2007.
31.Korn T, Oukka M, Kuchroo V, Bettelli E. Th17 cells: effector T cells with inflammatory properties. Semin Immunol 19(6):362-71, 2007.
32.Torchinsky MB, Blander JM. T helper 17 cells: discovery, function, and physiological trigger. Cell Mol Life Sci 67(9):1407-21, 2010.
33.Raza A, Yousaf W, Giannella R, Shata MT. Th17 cells: interactions with predisposing factors in the immunopathogenesis of inflammatory bowel disease. Expert Rev Clin Immunol 8(2):161-8, 2012.

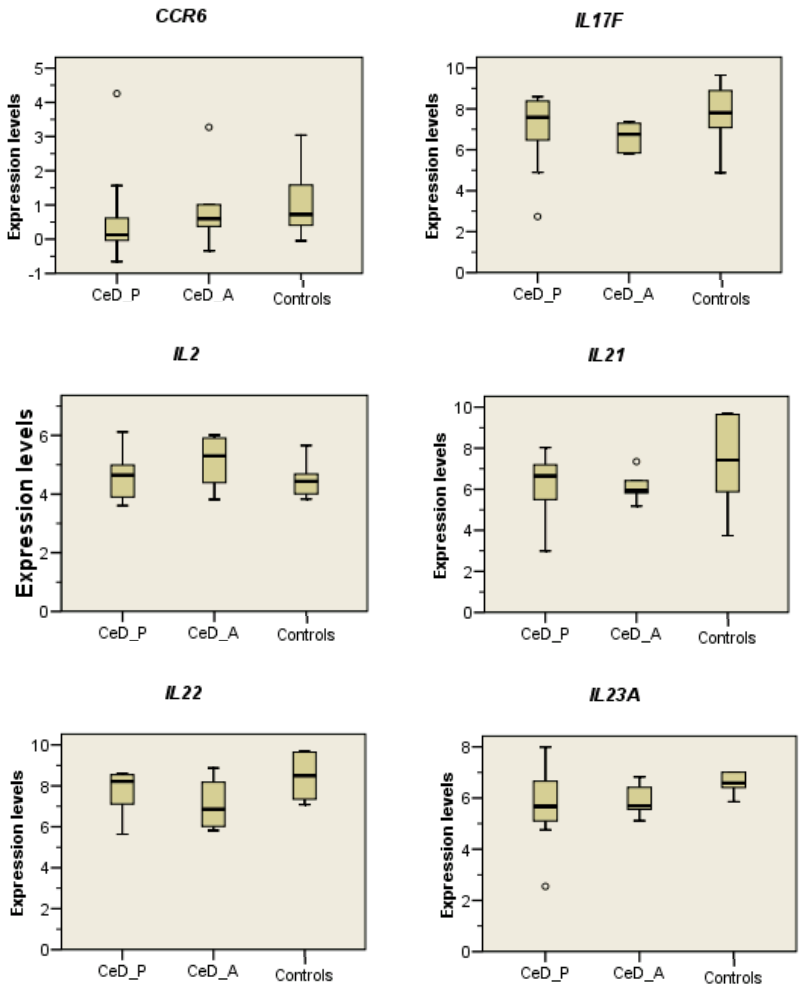



Figure 1. Expression levels (ΔCt) in adult (CeD_A) and pediatric CeD (CeD_P) celiac disease patients and in controls.*p<0.05, **p<0.01.

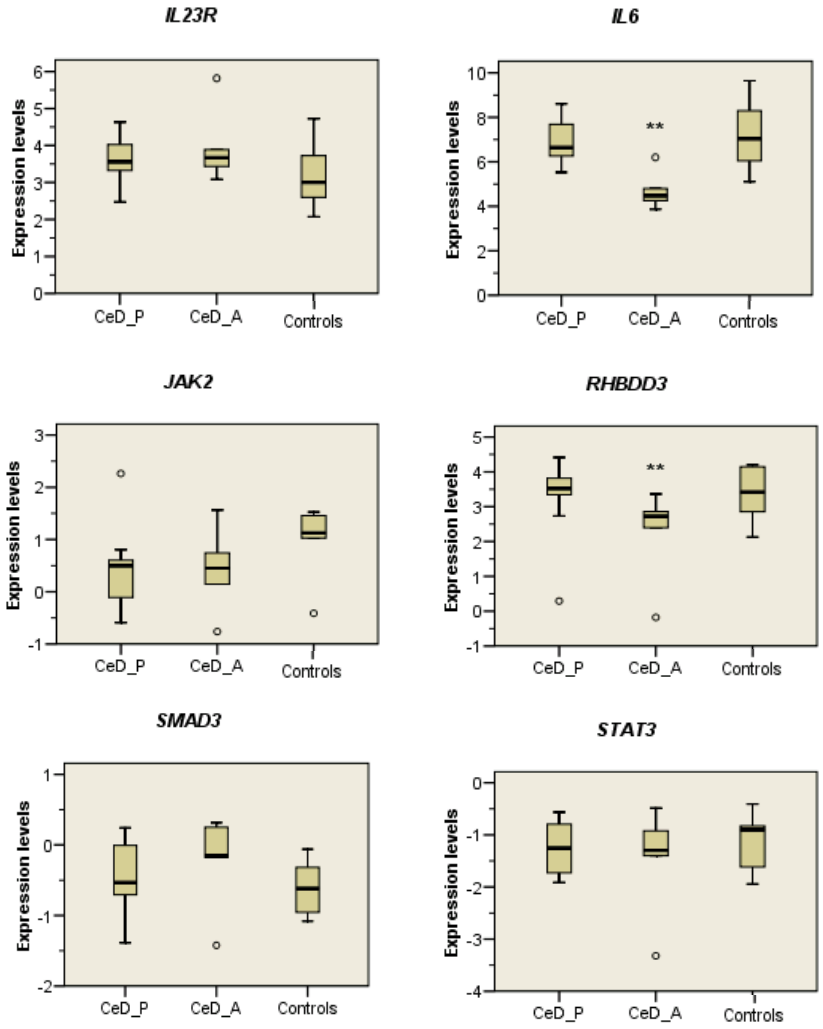



**Figure 1.** (continued) Expression levels (ΔCt) in adult (CeD_A) and pediatric CeD (CeD_P) celiac disease patients and in controls.*p<0.05, **p<0.01.

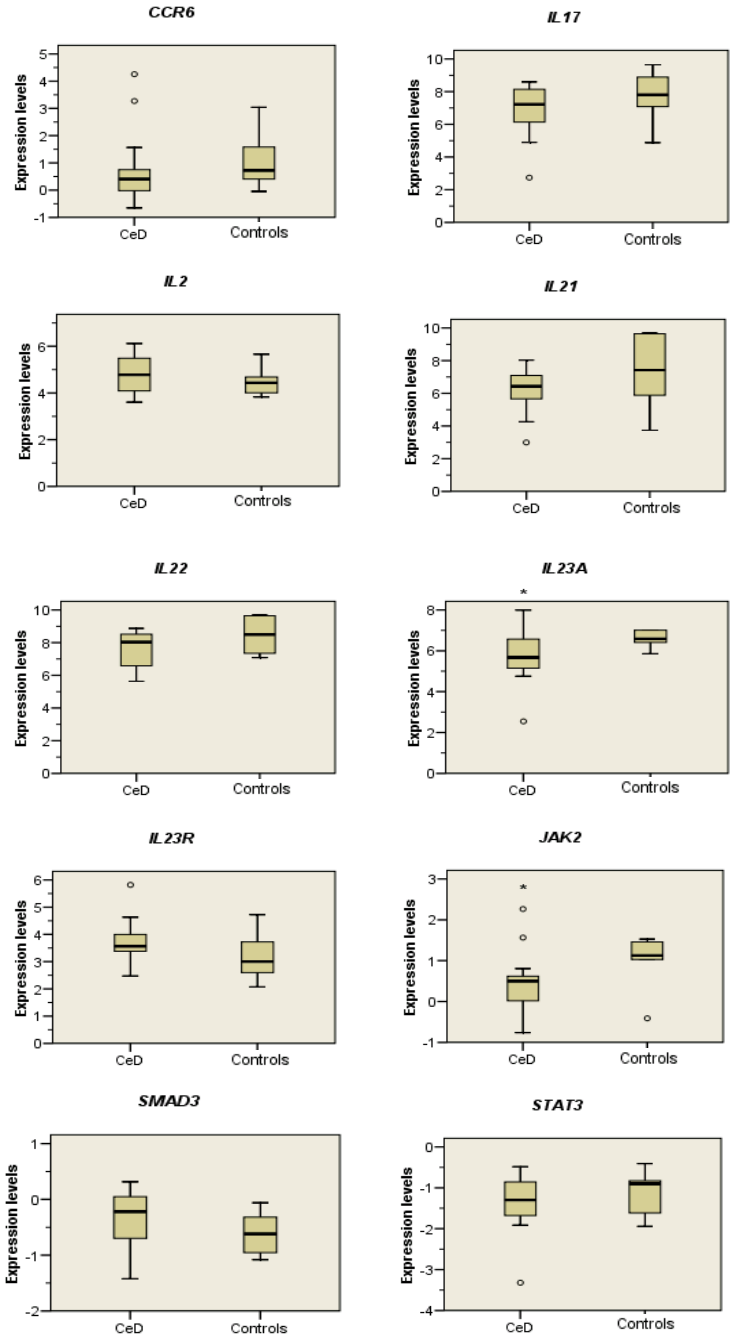



Figure 2. Expression levels (ΔCt) in celiac disease (CeD) patients and controls. *p<0.05, **p<0.01.

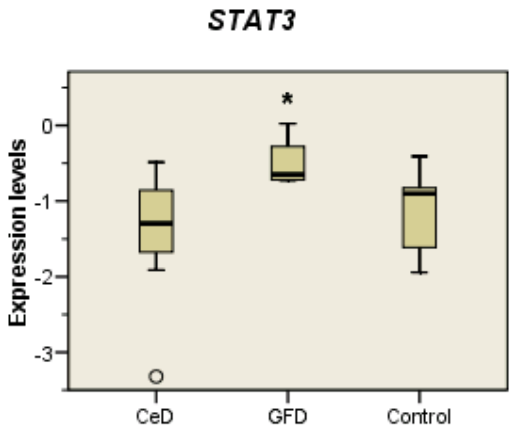



Figure 3. Expression level (ΔCt) of *STAT3* in celiac disease patients on gluten containing diet (CeD), celiac disease patients on a gluten free diet (GFD) and in controls;
*p<0.05.

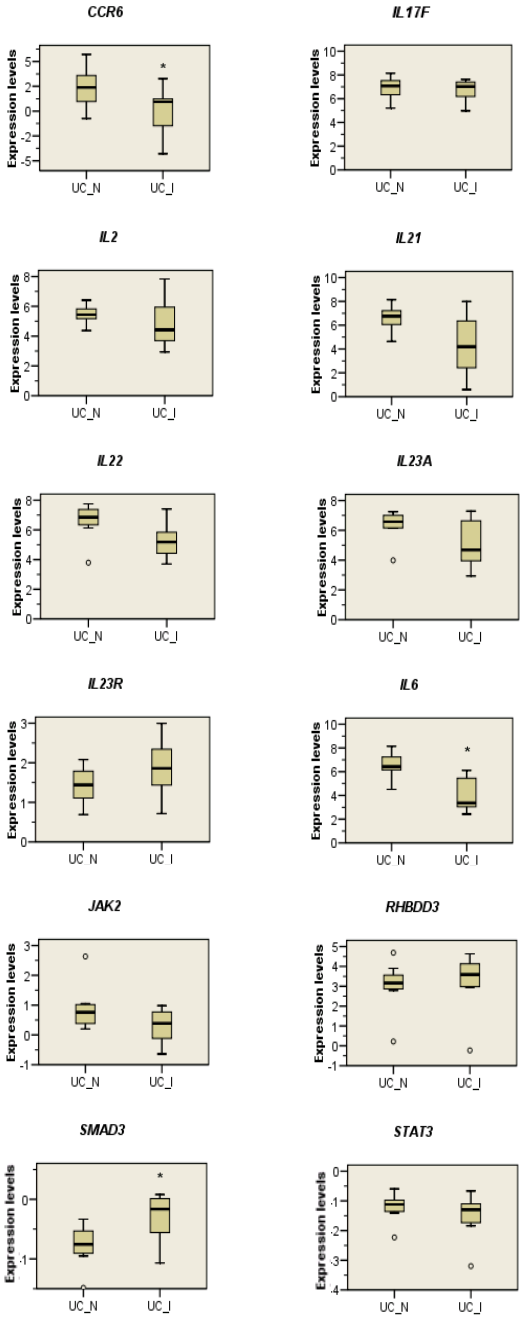



Figure 4. Expression levels (ΔCt) in normal UC (UC_N) vs. inflamed UC (UC_I); *p<0.05, **p<0.01.

# 3. ESTUDIO DE FARMACOGENÉTICA EN PACIENTES DE CROHN TRATADOS CON INFLIXIMAB

**I.** ***Role of TNFRSF1B polymorphisms in the response of Crohn´s disease patients to infliximab***

**II.** ***Validation of gene expression profile for response to infliximab in Crohn's disease***

**I.** ***Role of TNFRSF1B polymorphisms in the response of Crohn´s disease patients to infliximab***

L. M. Medrano[1] MSc, C. Taxonera[2] MD PhD, A Márquez[3] PhD, M. Barreiro-de Acosta[4] MD PhD, M. Gómez-García[5] MD PhD, C. González-Artacho[5] MD, J. L. Pérez-Calle[6] MD PhD, F. Bermejo[7] MD PhD, A. Lopez-Sanromán[8] MD PhD, M.D. Martín Arranz[9] MD PhD, J.P. Gisbert[10] MD PhD, J.L. Mendoza[2] MD PhD, J Martín[3] PhD, E Urcelay[1] PhD, and C. Núñez[1] PhD

[1] Immunology Department, Hospital Clínico San Carlos, and Instituto de Investigación Sanitaria del Hospital Clínico San Carlos (IdISSC), Madrid, Spain.
[2] Gastroenterology Department, Hospital Clínico San Carlos, and Instituto de Investigación Sanitaria del Hospital Clínico San Carlos (IdISSC), Madrid, Spain.
[3] Institute of Parasitology and Biomedicine, CSIC, Granada, Spain
[4] Department of Gastroenterology, Hospital Clínico Universitario de Santiago de Compostela, Spain.
[5] Department of Gastroenterology, Hospital Virgen de las Nieves, Granada, Spain
[6] Department of Gastroenterology, Hospital Alcorcón, Madrid, Spain.
[7] Department of Gastroenterology, Hospital Fuenlabrada, Madrid, Spain.
[8] Department of Gastroenterology, Hospital Ramón y Cajal, Madrid, Spain.
[9] Department of Gastroenterology, Hospital La Paz, Madrid, Spain.
[10] Department of Gastroenterology, Hospital Universitario de La Princesa, Instituto de Investigación Sanitaria Princesa (IP), Madrid, and Centro de Investigación Biomédica en Red de Enfermedades Hepáticas y Digestivas (CIBERHED), Spain.

**Hum Immunol. 2013 Oct 10. [Epub ahead of print]**

**Abstract**

Infliximab (IFX) is a valid treatment for Crohn´s disease (CD), but a relevant percentage of patients do not benefit from this therapy. In the Japanese population, the response to IFX was associated with markers in the TNF receptor superfamily 1A (*TNFRSF1A*) and 1B (*TNFRSF1B*) genes. We aimed to replicate the association previously described in the Japanese population and to ascertain the role of TNF receptors as modulators of the response to IFX. We studied 297 white Spanish CD patients with a known response to IFX: 238 responders and 59 primary nonresponders. Four single nucleotide polymorphisms (SNPs) were analysed: rs767455 in *TNFRSF1A* and rs1061622, rs1061624 and rs3397 in *TNFRSF1B*. Comparisons between groups were performed with chi-square tests or the Fisher´s exact test. Different features (sex, age, disease duration, smoking) were evaluated as possible confounding factors. No significant association was found between the studied *TNFRSF1A* polymorphisms and response to IFX. In the *TNFRSF1B* gene, the haplotype rs1061624_A-rs3397_T was significantly increased in nonresponders: p=0.015, OR=1.78, 95% CI 1.09-2.90; and an increased frequency of rs1061622_G carriers was observed in patients with remission: p=0.033 vs nonresponders and p=0.023 vs patients with a partial response. Our results support a role of *TNFRSF1B* gene variants in the response to IFX in CD patients.

**Key words:** Crohn´s disease, infliximab, pharmacogenetics, *TNFRSF1A*, *TNFRSF1B*.

**Introduction**

Crohn´s disease (CD) is characterized by a chronic inflammation of the gastrointestinal tract, which is most probably caused by an inappropriate immunological response to intestinal microbes in genetically susceptible individuals [1]. Recently, many genes have been associated with risk of disease, mainly due to the development of genome wide association studies (GWAS) [2]. Genetic factors also underlie the observed differences among individuals in response to the drugs used as therapy. This opens up the possibility of searching for genetic markers to assist us in the classification of patients as responders or nonresponders to specific treatments. Recent advances in CD treatments make such a classification particularly important. Ideally, the best treatment for each patient could be determined or, at least, the individual chance of success for a specific treatment could be predicted.

Tumour necrosis factor alpha (TNF-α) is a nonspecific mediator of tissue injury in CD. It can bind to two different receptors: TNFRSF1A (tumour necrosis factor receptor superfamily 1A) and TNFRSF1B (tumour necrosis factor receptor superfamily 1B); which activates NF-κB and triggers several inflammatory pathways [3]. Infliximab (IFX) is a chimeric monoclonal IgG1 that neutralizes TNF by binding it [4]. IFX has been proven to be effective for the treatment of both luminal [5] and fistulizing CD [6]. However, around 20-30% of CD patients do not show a good response to IFX [7]. In addition to the genetic factors underlying this variable response, other parameters such as previous surgery, concomitant treatments, smoking habits or disease duration may contribute to the observed variability.

The widespread use of IFX for CD treatment has led to an increase in the number of pharmacogenetic studies. The binding affinity and/or the kinetic stability of TNF for their receptors TNFRSF1A or TNFRSF1B could be affected by the presence of genetic variants in these receptors. In addition, polymorphisms in the genes coding these receptors could affect the signal transduction cascade. Therefore, TNF receptors are good candidates as response modifiers. Scarce information exist about the functional role of genetic polymorphisms in *TNFRSF1A* and *TNFRSF1B,* but it seems that the non-synonymous polymorphism rs1061622 and some haplotypes conformed by polymorphisms located in the 3´-untranslated region (UTR) (rs1061624, rs5030792 and rs3397), all related to the *TNFRSF1B* gene, may have functional consequences

through their role in signal transduction or in mRNA stability [8-10]. The role of some genetic polymorphisms in *TNFRSF1A*, located on chromosome 12p13, and *TNFRSF1B*, located on 1p36, in the response of CD patients to IFX have been previously studied, but consensus is lacking. In the Japanese population, genetic variants in *TNFRSF1A* and *TNFRSF1B*, including the previously described ones with a functional role, were studied and association with primary nonresponse to IFX was found in a sample of 80 patients, 32 of them classified as nonresponders using the Harvey-Bradshaw index (HBI) [11]. In Caucasian populations, apparently discrepant results have been reported [12-14] in studies with a variable sample size, but different criteria to evaluate response to IFX were followed (Crohn´s disease activity index, CDAI; overall clinical evaluation). Most of these studies are only focused on primary nonresponse to IFX. A secondary loss of response is also relevant, but it probably has a different etiology, as it is caused by different mechanisms, mainly pharmacodinamics mechanisms but also by immunogenicity and by non-immune-mediated pharmacokinetics [15].

We aimed to study the genetic polymorphisms in *TNFRSF1A* and *TNFRSF1B* previously associated with response to IFX in the Japanese population in our sample of IFX treated CD patients.

## Materials and Methods

Patients and criteria of response to IFX

We collected samples from CD patients treated with IFX from 7 centres around Spain: 238 responders and 59 primary nonresponders to this therapy. All of them were Spanish of Caucasian ancestry and at least 18 years old of age. All had been diagnosed with CD according to clinical, radiologic, endoscopic and histological criteria [16] and had received at least three induction doses of IFX (5 mg per kilogram) at weeks 0, 2 and 6. IFX was administered to treat either moderate to severe active luminal CD or active fistulizing perianal CD. All patients gave written informed consent and the study was approved by the ethic committees of all participating hospitals.

The response to IFX was determined by a chronological review of the medical records, and data were centrally monitored. In patients with luminal disease, response was evaluated by the Harvey-Bradshaw index (HBI) [17] at the beginning and 10 weeks after the first IFX dose. Partial response was defined as a decrease in the HBI

of more than 3 points and absence of concomitant corticosteroids. Remission was defined as a final HBI ≤4 and absence of concomitant corticosteroids. In patients with perianal disease, response was evaluated at week 10 after the first IFX dose. Remission was defined as the complete closure of all fistulas and partial response as a reduction (≥50%) in the number of draining fistulas. Patients who received IFX for both luminal and fistulizing disease and who achieved remission of any type (in the luminal or in the fistulizing disease) that justified maintenance treatment with IFX were considered as responders. All patients who did not achieve partial response or remission after the three IFX induction doses were considered as primary nonresponders.

Genotyping

We analysed four SNPs previously studied in the Japanese population within the *TNFRSF1A* and *TNFRSF1B* genes: rs767455 in a coding exon of *TNFRSF1A*; and rs1061622, rs1061624 and rs3397 in *TNFRSF1B*. Genotyping was performed by TaqMan technology (Applied Biosystems, Foster City, CA, USA) and at least 95% of samples were successfully genotyped for each SNP. Haplotypes were estimated using the EM (Expectation-Maximization) algorithm from the Haploview 4.1 software.

Statistical analyses

Genetic comparisons between responders and primary nonresponders were analysed by chi-square tests or the Fisher´s exact test (when the expected values were lower than five). Since we aimed to replicate previously described associations, no correction for multiple testing was necessary.

Comparisons between CD groups for other clinical or demographic characteristics corresponding to first IFX infusion (Table 1) were performed using chi-square tests for categorical variables and the Mann-Whitney U test for continuous variables. Those variables (sex, disease duration, smoking at first administration dose, concomitant immunomodulator treatment, etc) in addition to partial or total response to IFX were evaluated as possible confounding factors by logistic regression using the SPSS 15.0 software. Concomitant immunomodulator treatment was defined as the use of any immunosuppressive agent (azathioprine, mercaptopurine, methotrexate and mycophenolate mofetil) for at least 4 weeks before the first infliximab induction dose.

**Results**

All the polymorphisms included in the study conformed to Hardy-Weinberg expectations**.** Linkage disequilibrium between the SNPs in *TNFRSF1B* was as follows: D´=0.25 $r^2$=0.012 LOD=0.64 between rs1061622 and rs1061624; D´=0.10 $r^2$=0.001 LOD=0.08 between rs1061622 and rs3397; and D´=0.40 $r^2$=0.114 LOD=7.8 between rs1061624 and rs3397. Therefore, only haplotypes conformed by rs1061624 and rs3397, the two SNPs showing significant linkage disequilibrium, were considered.

Base-line characteristics

The comparison of several demographic and clinical characteristics between responders and nonresponders to IFX showed a significant difference for disease duration (Table 1).

Response vs nonresponse to IFX

The genetic frequencies for all the SNPs in responders and nonresponders are shown in Table 2. We did not observe significant differences between responders and nonresponders when studying the SNP located in the *TNFRSF1A* gene. In *TNFRSF1B*, the frequency of the allele A of rs1061624 was significantly higher in nonresponders (53.6%) than in responders to IFX (41.5%) (p=0.02; OR=1.63 95% CI 1.05-2.51). When studying rs3397, a statistically significant result was observed for the minor genotype (CC), which was increased in responders to IFX: p=0.05, OR=3.19 95% CI 0.95-16.78. The analysis of the haplotypes conformed by these two SNPs showed one haplotype, rs1061624_A rs3397_T, significantly increased in nonresponders (p=0.015, OR=1.78 95% CI 1.09-2.90) (Table 4).

Partial response to IFX vs remission

In Table 3, genotypic and allelic frequencies of the studied SNPs in patients showing partial response or remission after treatment with IFX are presented for the available patients. Significant differences are only observed when studying rs1061622. Comparison of each of those groups with nonresponders showed only significant differences when considering patients with remission: 29% carriers of the G allele in nonresponders vs 47% in patients showing remission (p=0.033) and vs 23% in patients with a partial response (p=0.55).

All the factors listed in Table 1 were tested as potential confounders of the association between the individual SNPs and the haplotypes and response to IFX, but none of them showed any effect.

Table 1. Characteristics of Crohn's disease patients at first infliximab infusion, both responders (n=238) and nonresponders (n=59) to infliximab.

| | Nonresponders | | Responders | | p value |
|---|---|---|---|---|---|
| | N | % | N | % | |
| **AGE** | 43.1 ±1.6 | | 39.9 ± 0.8 | | 0.619 |
| **SEX** | | | | | |
| MALE | 26 | 41.0 | 109 | 41.1 | 0.81 |
| FEMALE | 33 | 59.0 | 129 | 58.9 | |
| **DISEASE DURATION (years)** | 13.5 ± 1.1 | | 10.6 ± 0.5 | | 0.007 |
| **AGE AT DIAGNONOSIS (A)** | | | | | |
| A1 | 6 | 9.3 | 31 | 13.7 | |
| A2 | 45 | 79.1 | 175 | 72.6 | 0.83 |
| A3 | 7 | 11.6 | 29 | 13.7 | |
| **LOCATION (L)** | | | | | |
| L1 | 23 | 40.5 | 64 | 25.6 | 0.174[a] |
| L2 | 9 | 14.3 | 42 | 16.9 | |
| L3 | 25 | 45.0 | 117 | 55.6 | |
| L4 | 0 | 0 | 0 | 0 | |
| L1+L4 | 0 | 0 | 3 | 1.3 | |
| L2+L4 | 0 | 0 | 0 | 0 | |
| L3+L4 | 0 | 0 | 5 | 0.62 | |
| **BEHAVIOR (B)** | | | | | |
| B1 | 18 | 28.6 | 66 | 23.0 | 0.764 |
| B2 | 5 | 11.9 | 17 | 10.0 | |
| B3 | 9 | 16.7 | 33 | 12.5 | |
| B1p | 7 | 14.3 | 59 | 26.9 | |
| B2p | 2 | 0 | 10 | 5.0 | |
| B3p | 15 | 28.6 | 47 | 22.5 | |
| **SMOKING** | | | | | |
| YES | 20 | 31.8 | 105 | 47.7 | 0.076 |
| NO | 39 | 68.2 | 116 | 52.3 | |
| **IMMUNOMODULATOR TREATMENT[b]** | | | | | |
| YES | 44 | 93.5 | 172 | 92.9 | 0.464 |
| NO | 3 | 6.5 | 16 | 7.1 | |

A1: ≤16 years; A2: 17-40 years; A3: >40 years. L1: terminal ileum; L2: colon; L3: ileocolon; L4: Upper GI; L1+L4: terminal ileum+ upper GI; L2+L4: colon+upper GI; L3+L4: ileocolon+upper GI. B1: nonstricturing, nonpenetrating; B2: structuring; B3: penetrating; B1p: nonstricturing, nonpenetrating+perianal; B2p: stricturing+perianal; B3p: penetrating+perianal.

[a]Excluding categories with L4.

[b]Patients with no concomitant treatment have not been included in this comparison.

Table 2. Genotypic and allelic frequencies for the studied SNPs.

| | Nonresponders | | Responders | | p value | OR (95% CI) |
|---|---|---|---|---|---|---|
| | N | % | N | % | | |
| ***TNFRSF1A*** | | | | | | |
| rs767455 | | | | | | |
| AA | 25 | 45.5 | 99 | 43.4 | - | 1.0 (reference) |
| AG | 23 | 41.8 | 94 | 41.2 | 0.92 | 0.97 (0.49-1.91) |
| GG | 7 | 12.7 | 35 | 15.4 | 0.62 | 0.79 (0.28-2.15) |
| AG+GG | 30 | 54.5 | 129 | 56.6 | 0.78 | 0.92 (0.49-1.74) |
| A | 73 | 66.4 | 292 | 64.0 | 0.65 | 0.90 (0.57-1.43) |
| G | 37 | 33.6 | 164 | 36.0 | | |
| ***TNFRSF1B*** | | | | | | |
| rs1061622 | | | | | | |
| TT | 41 | 70.7 | 144 | 61.3 | - | 1.0 (reference) |
| TG | 16 | 27.6 | 84 | 35.7 | 0.21 | 0.67 (0.34-1.32) |
| GG | 1 | 1.7 | 7 | 3.0 | 0.51 | 0.50 (0.01-4.10) |
| TG+GG | 17 | 29.3 | 91 | 38.7 | 0.18 | 0.66 (0.34-1.27) |
| T | 98 | 84.5 | 372 | 79.1 | 0.20 | 0.70 (0.39-1.24) |
| G | 18 | 17.5 | 98 | 20.9 | | |
| rs1061624 | | | | | | |
| GG | 10 | 17.9 | 88 | 37.5 | - | 1.0 (reference) |
| GA | 32 | 57.1 | 99 | 42.1 | 0.006 | 2.84 (1.25-6.59) |
| AA | 14 | 25.0 | 48 | 20.4 | 0.033 | 2.57 (0.98-6.80) |
| GA+AA | 46 | 82.1 | 147 | 62.5 | 0.005 | 2.75 (1.26-6.15) |
| G | 52 | 46.4 | 275 | 58.5 | 0.021 | 1.63 (1.05-2.51) |
| A | 60 | 53.6 | 195 | 41.5 | | |
| rs3397 | | | | | | |
| CC | 3 | 5.2 | 35 | 14.8 | - | 1.0 (reference) |
| TC | 30 | 51.7 | 101 | 42.8 | 0.04 | 3.47 (0.98-18.73) |
| TT | 25 | 43.1 | 100 | 42.4 | 0.083 | 2.92 (0.81-15.93) |
| TT+TC | 55 | 94.8 | 201 | 85.2 | 0.05 | 3.19 (0.95-16.78) |
| T | 80 | 69.0 | 301 | 63.8 | 0.29 | 0.79 (0.50-1.25) |
| C | 36 | 31.0 | 171 | 36.2 | | |

The genotype indicated as OR=1.0 is used as the reference for the remaining genotypes and for the carrier to calculate p values and OR.

Allelic ORs are relative to the minor allele in responders.

Table 3. Genotypic and allelic frequencies for the studied SNPs in patients stratified by type of response.

| | Partial response | | Remission | | p value | OR (95% CI) |
|---|---|---|---|---|---|---|
| | N | % | N | % | | |
| ***TNFRSF1A*** | | | | | | |
| rs767455 | | | | | | |
| AA | 14 | 45.2 | 40 | 40.8 | - | 1.0 (reference) |
| AG | 13 | 41.9 | 39 | 39.8 | 0.91 | 0.95 (0.36-2.49) |
| GG | 4 | 12.9 | 19 | 19.4 | 0.42 | 0.60 (0.13-2.29) |
| AG+GG | 17 | 54.8 | 58 | 59.2 | 0.67 | 0.84 (0.34-2.04) |
| A | 41 | 66.1 | 119 | 60.7 | 0.44 | 0.79 (0.42-1.50) |
| G | 21 | 33.9 | 77 | 39.3 | | |
| ***TNFRSF1B*** | | | | | | |
| rs1061622 | | | | | | |
| TT | 23 | 76.7 | 54 | 53.5 | - | 1.0 (reference) |
| TG | 6 | 20.0 | 45 | 44.6 | 0.016 | 0.31 (0.10-0.89) |
| GG | 1 | 3.3 | 2 | 2.0 | 0.898 | 1.17 (0.02-23.56) |
| TG+GG | 7 | 23.3 | 47 | 46.6 | 0.023 | 0.35 (0.12-0.96) |
| T | 52 | 86.7 | 153 | 75.7 | 0.071 | 0.48 (0.20-1.14) |
| G | 8 | 13.3 | 49 | 24.3 | | |
| rs1061624 | | | | | | |
| GG | 11 | 35.5 | 37 | 36.3 | - | 1.0 (reference) |
| GA | 11 | 35.5 | 41 | 40.2 | 0.83 | 0.90 (0.32-2.57) |
| AA | 9 | 29.0 | 24 | 23.5 | 0.66 | 1.26 (0.40-3.92) |
| GA+AA | 19 | 64.5 | 65 | 63.7 | 0.97 | 0.98 (0.39-2.49) |
| G | 33 | 53.2 | 115 | 56.4 | 0.66 | 1.14 (0.62-2.09) |
| A | 29 | 46.8 | 89 | 43.6 | | |
| rs3397 | | | | | | |
| CC | 4 | 12.9 | 18 | 17.6 | - | 1.0 (reference) |
| TC | 11 | 35.5 | 44 | 43.1 | 0.86 | 1.13 (0.28-5.48) |
| TT | 16 | 51.6 | 40 | 39.2 | 0.34 | 1.80 (0.48-8.39) |
| TT+TC | 27 | 87.1 | 84 | 82.3 | 0.53 | 1.45 (0.42-6.37) |
| T | 43 | 69.4 | 124 | 60.8 | 0.22 | 0.68 (0.36-1.31) |
| C | 19 | 30.6 | 80 | 39.2 | | |

Partial response and remission data are shown for the available patients.

The genotype indicated as OR=1.0 is used as the reference for the remaining genotypes and for the carrier to calculate p values and OR.

Allelic ORs are relative to the minor allele in responders.

Table 4. Frequency of the haplotypes conformed by the studied SNPS in *TNFRSF1B* in linkage disequilibrium (rs1061624, rs3397) in nonresponders (2N=118) and in responders (2N=476).

| Haplotype | Nonresponders (%) | Responders (%) | p value | OR (95% CI) |
|---|---|---|---|---|
| GT | 39.6 | 45.5 | 0.2496 | |
| AC | 24.6 | 23.2 | 0.7377 | |
| AT | 28.5 | 18.3 | 0.0142 | 1.78 (1.09-2.90) |
| GC | 7.3 | 13.1 | 0.0869 | |

**Discussion**

We studied the association of several genetic polymorphisms located in *TNFRSF1A* and *TNFRSF1B* with response to IFX in CD patients. When considering several demographic and clinical characteristics, a better response to IFX was observed in patients with an earlier treatment, as previously described [18]. In addition, we found that one haplotype conformed by two SNPs studied in *TNFRSF1B* (rs1061624_A-rs3397_T) was significantly increased in nonresponders to this therapy. A similar result was previously observed in the Japanese population following the same criteria of response to IFX that the one used in this work [11], which suggest the existence of a genetic variant in linkage disequilibrium with this haplotype modulating the response to IFX. We also found a genetic variant, rs1061622, significantly increased in patients showing remission after IFX treatment. Steenholdt et al [14] also reported an increased frequency of carriage of rs1061622_C in this particular group of patients. It is very interesting that a functional role has been described for one haplotype containing the rs1061624 and rs3397 polymorphisms [9] and for the variant rs1061622 .

The role of genetic variants in the genes coding for TNF receptors as predictors of response to IFX for CD treatment has been previously studied by several groups [11-14]. It should be highlighted that all those studies included a reduced number of SNPs and therefore no conclusive results could be obtained. However, lack of association between *TNFRSF1A* polymorphisms and response to IFX has been quite consistently found, which dampened interest in this gene. The case of *TNFRSF1B* is different, with several polymorphisms associated with response to IFX by different groups. The haplotype rs1061624_A-rs3397_T was initially associated

with IFX response in the Japanese population and no data exist in Caucasians. Mascheretti et al [12] studied those two SNPs in the German population but haplotype analyses were not performed. These authors did not find an association after studying the individual SNPs; such association was not detected in the Japanese study either. It should be noted that the distribution of these alleles in their carrying haplotypes will influence the chance of obtaining a significant result, which will be obtained only when one allele is mainly present in the associated haplotype.

A high rate of false positive results has been suggested as responsible for the discrepant results observed in pharmacogenetic studies, which would be influenced by the inclusion of low numbers of patients. However, replication of a significant result in a different sample or population is warranted to validate a reported association and it avoids the necessity of correction for multiple testing. Thus, after our significant result, *TNFRSF1B* seems to contain a genetic variant involved in the clinical response to IFX. A role of this gene is also supported by the reported association of *TNFRSF1B* variants with biological response to IFX [13, 14] and by the differences in *TNFRSF1B* expression observed depending on how the patients respond to this therapy [14, 19]. These findings support the influence of this gene in the response of CD patients to IFX treatment.

It is also important to keep in mind that the apparent discrepancy between studies should be interpreted with caution because of differences in the criteria for evaluating response to IFX.

A limitation of the study could be a potential bias in the assessment of response due to the lack of endoscopy or MRI. In clinical trials the definition of response or remission has been based on clinical activity indexes and not on CRP values or endoscopic or MRI assessment [5,6].

The ultimate objective of pharmacogenetics, application of a personalized medicine, remains a distant goal given the low OR observed. Nevertheless, although the haplotype rs1061624_A-rs3397_T does not represent a useful pharmacogenetic marker, the situation might be different when the functional variant modulating the response to IFX is identified. Much more genetic variation than the one studied in *TNFRS1B* exists and a higher effect is expected for the functional variant. Unfortunately, the sample size of our study does not provide with enough statistical power to look for this variant. A large undertaking, most probably involving international collaboration, is now expected to attain this objective. The approach

would involve collecting samples from a large number of CD patients with homogeneous criteria for IFX administration and for evaluation of the clinical response.

**Acknowledgments**

We thank patients and controls for making this study feasible. Angel García and Carmen Martínez provided expert technical assistance.

**References**

1. Abraham C, Cho JH (2009) Inflammatory bowel disease. N Engl J Med 361: 2066-2078.

2. Franke A, McGovern DP, Barrett JC, Wang K, Radford-Smith GL, et al. (2010) Genome-wide meta-analysis increases to 71 the number of confirmed Crohn's disease susceptibility loci. Nat Genet 42: 1118-1125.

3. Monteleone G, Pallone F, MacDonald TT (2011) Emerging immunological targets in inflammatory bowel disease. Curr Opin Pharmacol 11: 640-645.

4. Van Assche G, Rutgeerts P (2000) Anti-TNF agents in Crohn's disease. Expert Opin Investig Drugs 9: 103-111.

5. Rutgeerts P, Feagan BG, Lichtenstein GR, Mayer LF, Schreiber S, et al. (2004) Comparison of scheduled and episodic treatment strategies of infliximab in Crohn's disease. Gastroenterology 126: 402-413.

6. Sands BE, Anderson FH, Bernstein CN, Chey WY, Feagan BG, et al. (2004) Infliximab maintenance therapy for fistulizing Crohn's disease. N Engl J Med 350: 876-885.

7. Chaudhary R, Ghosh S (2005) Prediction of response to infliximab in Crohn's disease. Dig Liver Dis 37: 559-563.

8. Morita C, Horiuchi T, Tsukamoto H, Hatta N, Kikuchi Y, et al. (2001) Association of tumor necrosis factor receptor type II polymorphism 196R with Systemic lupus erythematosus in the Japanese: molecular and functional analysis. Arthritis Rheum 44: 2819-2827.

9. Puga I, Lainez B, Fernandez-Real JM, Buxade M, Broch M, et al. (2005) A polymorphism in the 3' untranslated region of the gene for tumor necrosis factor receptor 2 modulates reporter gene expression. Endocrinology 146: 2210-2220.

10. Till A, Rosenstiel P, Krippner-Heidenreich A, Mascheretti-Croucher S, Croucher PJ, et al. (2005) The Met-196 -> Arg variation of human tumor necrosis factor receptor 2 (TNFR2) affects TNF-alpha-induced apoptosis by impaired NF-kappaB signaling and target gene expression. J Biol Chem 280: 5994-6004.
11. Matsukura H, Ikeda S, Yoshimura N, Takazoe M, Muramatsu M (2008) Genetic polymorphisms of tumour necrosis factor receptor superfamily 1A and 1B affect responses to infliximab in Japanese patients with Crohn's disease. Aliment Pharmacol Ther 27: 765-770.
12. Mascheretti S, Hampe J, Kuhbacher T, Herfarth H, Krawczak M, et al. (2002) Pharmacogenetic investigation of the TNF/TNF-receptor system in patients with chronic active Crohn's disease treated with infliximab. Pharmacogenomics J 2: 127-136.
13. Pierik M, Vermeire S, Steen KV, Joossens S, Claessens G, et al. (2004) Tumour necrosis factor-alpha receptor 1 and 2 polymorphisms in inflammatory bowel disease and their association with response to infliximab. Aliment Pharmacol Ther 20: 303-310.
14. Steenholdt C, Enevold C, Ainsworth MA, Brynskov J, Thomsen OO, et al. (2012) Genetic polymorphisms of tumour necrosis factor receptor superfamily 1b and fas ligand are associated with clinical efficacy and/or acute severe infusion reactions to infliximab in Crohn's disease. Aliment Pharmacol Ther 36: 650-659.
15. Steenholdt C, Brynskov J, Thomsen OO, Munck LK, Fallingborg J, et al. (2013) Individualised therapy is more cost-effective than dose intensification in patients with Crohn's disease who lose response to anti-TNF treatment: a randomised, controlled trial. Gut.
16. Lennard-Jones JE (1989) Classification of inflammatory bowel disease. Scand J Gastroenterol Suppl 170: 2-6; discussion 16-19.
17. Harvey RF, Bradshaw JM (1980) A simple index of Crohn's-disease activity. Lancet 1: 514.
18. Miheller P, Lakatos PL, Horvath G, Molnar T, Szamosi T, et al. (2009) Efficacy and safety of infliximab induction therapy in Crohn's Disease in Central Europe--a Hungarian nationwide observational study. BMC Gastroenterol 9: 66.
19. Arijs I, Quintens R, Van Lommel L, Van Steen K, De Hertogh G, et al. (2010) Predictive value of epithelial gene expression profiles for response to infliximab in Crohn's disease. Inflamm Bowel Dis 16: 2090-2098.

**Conflict of interest**

Authors' declaration of personal interests: C. Taxonera, M. Barreiro-de Acosta, A. Lopez-Sanroman, M.D. Martín Arranz, J.P. Gisbert, J.L. Mendoza have served as speakers, consultants and advisory members for MSD and Abbott.

Declaration of funding interests: this work was funded by Fondo de Investigaciones Sanitarias FIS (PI11/00614) and Fundación Mutua Madrileña.

II. ***Validation of gene expression profile for response to infliximab in Crohn's disease***

L. M. Medrano[1*] MSc, C. Taxonera[2*] MD PhD, M. Barreiro-de Acosta[3] MD PhD, J. L. Pérez-Calle[4] MD PhD, F. Bermejo[5] MD PhD, A. Lopez-Sanromán[6] MD PhD, M.D. Martín Arranz[7] MD PhD, J.P. Gisbert[8] MD PhD, J.L. Mendoza[2] MD PhD, M. Díaz-Rubio[2] MD PhD, E. G. de la Concha[1] MD PhD, C. Núñez[1] PhD, and E. Urcelay[1] PhD

[1] Immunology and [2] Gastroenterology Departments, Hospital Clínico San Carlos, Instituto de Investigación Sanitaria del Hospital Clínico San Carlos (IdISSC), Madrid, Spain.
[3] Department of Gastroenterology, Hospital Clínico Universitario de Santiago de Compostela, Spain.
[4] Department of Gastroenterology, Hospital Alcorcón, Madrid, Spain.
[5] Department of Gastroenterology, Hospital Fuenlabrada, Madrid, Spain.
[6] Department of Gastroenterology, Hospital Ramón y Cajal, Madrid, Spain.
[7] Department of Gastroenterology, Hospital La Paz, Madrid, Spain.
[8] Department of Gastroenterology, Hospital Universitario La Princesa, Instituto de Investigación Sanitaria Princesa (IP), Madrid, and Centro de Investigación Biomédica en Red de Enfermedades Hepáticas y Digestivas (CIBERHED), Spain.

[*] These authors contributed equally to this work

**(Enviado)**

**Summary**

Substantial proportion of Crohn's disease (CD) patients shows no response or only a limited response to treatment with infliximab (IFX). The identification of predictors of response would be of great clinical and economic benefit, allowing the targeting of this therapy to the patients who are most likely to respond. The expression profile of five genes (*TNFAIP6*, *G0S2*, *IL11*, *S100A8* and *S100A9)* was reported to predict response to IFX with an overall accuracy of 100%.

In an independent Spanish cohort, we aimed at validating this expression profile through genetic association analysis of these five genes. The study of genetic polymorphisms which could be underlying the observed differential expression would help to get more insight into the causal mechanism related with response to IFX and could lead to establish an easier classification of CD patients with regard to response to this treatment.

Patients (210) with active CD who received at least three induction doses of IFX were included and subsequently classified according to IFX response. A tagging strategy was used to select genetic polymorphisms that cover the variability present in the chromosomal regions where the tested genes with altered expression were located. Several demographic and clinical characteristics were analyzed as potential confounding factors of the IFX response using logistic regression.

Following genotyping, significant differences between responders and nonresponders to IFX were observed in haplotypes of two genetic regions: *S100A8-S100A9* (rs11205276*G/rs3014866*C/rs724781*C/rs3006488*A) and *IL11* (rs1126760*C/rs1042506*T). None of the characteristics tested as potential confounders was found to cause any effect. Moreover, combined analysis of the 12-tagging SNPs rendered a specific combination including the significant *IL11* and *S100* haplotypes, with frequencies of 1.3% in responders and 7.8% in nonresponders.

Our results successfully validate the reported expression signature predictive of anti-TNF outcome in CD patients and suggest an etiological role of those top-five genes in the IFX response pathway.

**Key Words**: *TNFAIP6*, *G0S2*, *IL11*, *S100A8, S100A9*, Calprotectin, Crohn's disease, pharmacogenetics.

**Introduction**

Crohn's disease (CD) is one of the clinical forms of inflammatory bowel disease (IBD) resulting from a defective regulation of mucosal immune responses to commensal microbiota in genetically susceptible individuals.[1] The last years have contemplated substantial progress in the identification of the genes involved in CD predisposition, propelled by the HapMap project and genome-wide association studies.[2] A better understanding of the biological pathways underlying CD pathogenesis will lead to the development of new therapeutic approaches that specifically target those pathways, and will eventually allow personalized treatments. Therefore, an increasing need exists to predict the therapy most fitted to each patient.

Since 1998, when the U.S. Food and Drug Administration approved infliximab (IFX) for treatment of moderate or severe CD that does not respond to conservative treatment, monoclonal antibodies to tumor necrosis factor alpha (TNF-□) have become the hallmark treatment for refractory CD. Infliximab has proven to be effective for the treatment of both luminal [3] and fistulizing CD. [4] However, a lack of response or a partial response to IFX has been consistently observed and a growing need exists to identify predictors of response in order to achieve a more efficacious use of this expensive and potentially toxic therapy. On the other hand, data from clinical trials of IFX suggest that high-risk patients and patients with active inflammation may benefit from earlier use of this drug. [5] Clinical parameters such as concurrent therapies, smoking habits or previous surgery seem to account for only a small amount of the variance in response to anti-TNF therapies. [6]

In a recent study, pretreatment mucosal gene expression profiles predicted response to first IFX treatment in CD patients by microarray analysis.[7] The study identified a 100% accurate predictive gene signature for response to IFX in Crohn's colitis; class prediction analysis allowed complete separation between responders and nonresponders through a panel of 5 top significant genes: *TNFAIP6*, *IL11*, *G0S2*, *S100A8* and *S100A9*. Despite these promising results, to date several studies used genome-wide expression analysis to identify expression signatures predicting response to anti-TNF treatment in rheumatoid arthritis patients, but results showed little overlap.[8] The expression profiles identified in different studies were often not consistent with each other and different gene sets were reported to distinguish between responders and nonresponders. Therefore, the expression signature found in CD warrants validation and it could be considered a first step towards a diagnostic test.

Provided corroboration is achieved, the cluster of five genes with decreased expression in CD responders compared to nonresponders to IFX therapy would be a useful tool to classify patients. With this hypothesis, we aimed at validating the reported expression profile by exploring the association of variants in those top-five genes with response to IFX in an independent Spanish cohort of CD patients. Once an association between polymorphisms within those genes and the response to IFX is found, a causal mechanism of IFX response would be envisaged and a simpler way to classify patients could be established.

## Material and Methods

### Ethics Statement

The Ethics Committees of the participant hospitals (Comité Etico de Investigación Clínica, CEIC, of the following hospitals from Madrid: Clínico San Carlos, Fuenlabrada, Alcorcón, La Princesa, La Paz and Ramón y Cajal, and CEIC of Galicia) approved the study and all patients provided written informed consent. The informed consent was approved by the Ethics Committee of the leading center (CEIC of Hospital Clínico San Carlos, Madrid).

### Study design and Patients

Overall, 210 unrelated white Spanish patients with active CD were consecutively recruited from 7 centers. Eligible patients were at least 18 years of age, had an established diagnosis of CD and had received at least the 3 induction doses of IFX (5 mg per kilogram) at weeks 0, 2 and 6. Diagnosis of CD was based on standard clinical, radiologic, endoscopic and histological criteria.[9] IFX was administered to treat either moderate to severe active luminal CD or active fistulizing perianal CD.

Demographic data collected included sex, age, smoking status, duration of disease, extraintestinal manifestations and concurrent use of corticosteroids, immunomodulators or mesalazine. Disease phenotype was determined following the Montreal Classification: age at diagnosis (A1, ≤16 years; A2, 17-40 years; A3, >40 years), anatomic location (L1, terminal ileum; L2, colon; L3, ileocolon; L4, upper gastrointestinal tract; and +L4, upper gastrointestinal modifier) and disease behavior (B1, inflammatory; B2, stricturing; B3, penetrating; and p, perianal modifier)[10].

Patients were classified as responders (remission or partial response) or nonresponders to IFX. The response to IFX was determined by a chronological review of the medical records, and data were centrally monitored. The response to IFX in patients with luminal disease was evaluated by the Harvey-Bradshaw index (HBI) [11] at the beginning and 10 weeks after the first IFX dose. Partial response was defined as a decrease in the HBI of more than 3 points and absence of concomitant corticosteroids.[12] Remission was defined as a final HBI ≤4 and absence of concomitant corticosteroids.[12] In patients with perianal disease, response was evaluated at week 10 after the first IFX dose. Remission was defined as the complete

closure of all fistulas and partial response as a reduction (≥50%) in the number of draining fistulas. Patients either receiving IFX for both luminal and fistulizing disease or achieving remission of any type that justified the maintenance treatment with IFX, were considered responders. All patients who did not achieve partial response or remission after the 3 IFX induction doses were considered nonresponders.

**Genotyping**

In order to cover the highest variability within each gene showing altered expression, we chose single nucleotide polymorphisms (SNPs) by aggressive tagging from the HapMap B36 CEU population, which captured markers with $r^2 > 0.8$ (mean $r^2 = 0.93$) and a minor-allele frequency (MAF)> 0.1. Genotyping of the Spanish samples was carried out with pre-designed TaqMan Assays from Applied Biosystems (Applied Biosystems Inc., Foster City, CA, USA), using 384 well plates in a 7900HT Fast Real-Time PCR system, under conditions recommended by the manufacturer. Genotyping call-rate success was over 95% for the SNPs in all groups of patients.

**Data analysis**

Demographic and baseline characteristics were compared between responders and nonresponders by using the Mann-Withney U test or the chi square test, whether continuous or categorical variables were considered.

The statistical analysis to compare allelic and genotypic distributions was performed using chi-square test or Fisher's exact test (when expected values were below 5). Odds ratios (ORs) were calculated and their 95% confidence intervals were estimated using the Cornfield method. Haplotypic frequencies were inferred with the Expectation-Maximization algorithm implemented in the Haploview 4.1 software. Linkage disequilibrium was measured by calculating two parameters, $r^2$ and D' (Fig 1). Demographic or clinical characteristics and concomitant treatments were analyzed as potential confounding factors of the IFX response using logistic regression.

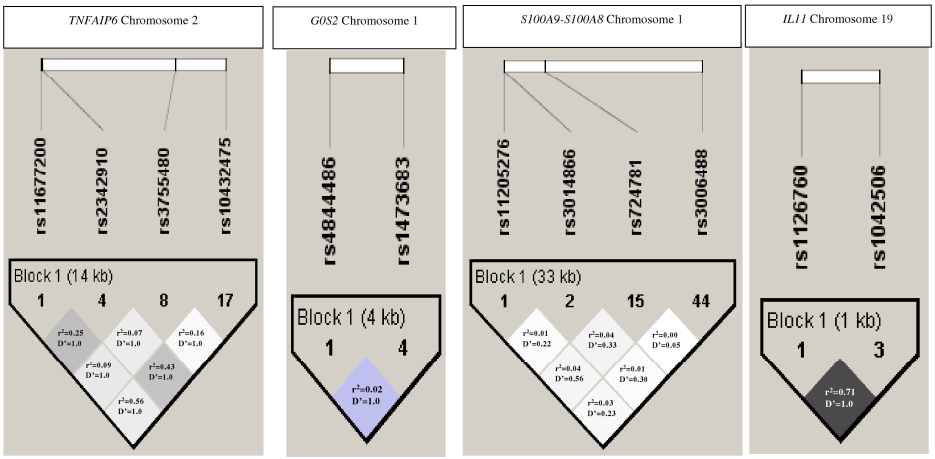



Figure 1. D' and $r^2$ of studied genes

**Results**

Baseline characteristics of the Spanish patients, classified as responders or nonresponders to IFX therapy, are summarized in Table 1. A total of 166 (79%) patients were classified as responders and 44 (21%) as nonresponders. No statistically significant differences were observed between both groups, other than a trend for smoking habit. In our cohort of Spanish patients, the use of any immunomodulator at baseline was not associated with response to IFX (93% of responders vs. 94% of nonresponders received some immunomodulator at baseline). Neither concomitant use of immunomodulators nor any of the other factors analyzed (presence of extraintestinal manifestations, familial history, indication criteria for IFX (luminal, perianal or both) and perianal consideration as a modifier in the Montreal Classification) showed a significant difference between both groups based on IFX response.

**Table 1.** Characteristics of Crohn's disease patients studied, both responders (n=166) and nonresponders (n=44) to infliximab

| | **RESPONDERS** | | **NONRESPONDERS** | | **P value** |
|---|---|---|---|---|---|
| | **N** | **%** | **N** | **%** | |
| **AGE** | 40.3 ± 0.9 | | 42.1 ±1.9 | | 0.37 |
| **SEX** | | | | | |
| MALE | 67 | 41.1 | 18 | 41.0 | 0.98 |
| FEMALE | 96 | 58.9 | 26 | 59.0 | |
| **YEARS OF DISEASE** | 12.5 ± 0.6 | | 14.1±1.29 | | 0.22 |
| **AGE AT DIAGNONOSIS (A)** | | | | | |
| A1 | 22 | 13.7 | 4 | 9.3 | |
| A2 | 117 | 72.6 | 34 | 79.1 | 0.67 |
| A3 | 22 | 13.7 | 5 | 11.6 | |
| **LOCATION (L)** | | | | | |
| L1 | 41 | 25.6 | 17 | 40.5 | 0.19[a] |
| L2 | 27 | 16.9 | 6 | 14.3 | |
| L3 | 89 | 55.6 | 19 | 45.0 | |
| L4 | 0 | 0 | 0 | 0 | |
| L1+L4 | 2 | 1.3 | 0 | 0 | |
| L2+L4 | 0 | 0 | 0 | 0 | |
| L3+L4 | 1 | 0.62 | 0 | 0 | |

**Table 1.** (continued)

| | RESPONDERS | | NONRESPONDERS | | P value |
|---|---|---|---|---|---|
| | **N** | **%** | **N** | **%** | |
| **BEHAVIOR (B)** | | | | | |
| B1 | 37 | 23.0 | 12 | 28.6 | 0.33 |
| B2 | 16 | 10.0 | 5 | 11.9 | |
| B3 | 20 | 12.5 | 7 | 16.7 | |
| B1p | 43 | 26.9 | 6 | 14.3 | |
| B2p | 8 | 5.0 | 0 | 0 | |
| B3p | 36 | 22.5 | 12 | 28.6 | |
| **SMOKING** | | | | | |
| YES | 71 | 47.7 | 14 | 31.8 | 0.06 |
| NO | 78 | 52.3 | 30 | 68.2 | |
| **IMMUNOMODULATOR TREATMENT[b]** | | | | | |
| YES | 105 | 92.9 | 29 | 93.5 | 1.00 |
| NO | 8 | 7.1 | 2 | 6.5 | |

Data correspond at first IFX.

A1: ≤16 years; A2: 17-40 years; A3: >40 years. L1: terminal ileum; L2: colon; L3: ileocolon; L4: Upper GI; L1+L4: terminal ileum+ upper GI; L2+L4: colon+upper GI; L3+L4: ileocolon+upper GI. B1: nonstricturing, nonpenetrating; B2: structuring; B3: penetrating; B1p: nonstricturing, nonpenetrating+perianal; B2p: structuring+perianal: B3p: penetrating+perianal.

[a]Excluding categories with L4.

[b]Patients with no concomitant treatment have not been included in this comparison.

Genotypic frequencies of these polymorphisms in the two groups of CD patients, responders and nonresponders to IFX therapy, are shown in Table 2. The studied polymorphisms conformed to Hardy-Weinberg expectations. Only rs3014866 in the *S100A9* gene showed a significant result. The aggressive tagging approach allowed a more thorough scrutiny of the studied regions by analyzing the haplotypes conformed by those polymorphisms (haplotypic frequencies in responders and non-

responders are summarized in Table 3). In the region where the *S100A9, S100A12* and *S100A8* genes map, the most frequent haplotype evidenced a significantly higher frequency in CD non-responders than in responders to IFX [p=0.02; OR (95% CI)= 1.81 (1.04- 3.12)]. Additionally, one haplotype in the *IL11* region showed a marginal significant association with the response to the anti-TNF treatment [p=0.06; OR (95% CI)= 2.03 (0.89- 4.59)]. No significant differences between the two groups of CD patients could be observed for haplotypes within the other two chromosomal regions explored. Nonetheless, when the studied polymorphisms were considered altogether (Table 4), the specific combination harboring the significant *IL11* and *S100* haplotypes showed a frequency of 1.3% in responders and of 7.8% in nonresponders and yielded an OR (95% CI)= 6.39 (1.32- 32.98). This combination seems to evidence a stronger effect than both significant *IL11* and *S100* haplotypes together which rendered an OR (95% CI)= 5 (0.65- 38.03) and it is, therefore, indicative of the coordinated effect of those 4 genomic regions.

The factors listed in Table 1 were tested as potential confounders in the analysis of all the individual SNPs and the haplotypes associated with the response to IFX therapy, but none of them showed an effect.

Table 2. Genotype frequencies of the polymorphisms located in the genes studied in Crohn's disease patients, both responders and nonresponders to infliximab

| | **RESPONDERS** | | **NONRESPONDERS** | |
|---|---|---|---|---|
| | **N** | **%** | **N** | **%** |
| **GENE *TNFAIP6*** | | | | |
| **rs11677200** | | | | |
| TT | 45 | 29 | 16 | 41 |
| TC | 89 | 57 | 15 | 38 |
| CC | 23 | 15 | 8 | 21 |
| **rs2342910** | | | | |
| AA | 86 | 56 | 21 | 60 |
| AT | 58 | 38 | 11 | 31 |
| TT | 10 | 6 | 3 | 9 |
| **rs3755480** | | | | |
| GG | 124 | 77 | 29 | 73 |
| GA | 33 | 20 | 9 | 23 |
| AA | 4 | 2 | 2 | 5 |
| **rs10432475** | | | | |
| AA | 130 | 82 | 32 | 84 |
| AG | 29 | 18 | 6 | 16 |
| GG | 0 | 0 | 0 | 0 |
| **GENE *G0S2*** | | | | |
| **rs4844486** | | | | |
| CC | 90 | 56 | 20 | 50 |
| CA | 63 | 39 | 19 | 48 |
| AA | 9 | 6 | 1 | 3 |
| **rs1473683** | | | | |
| GG | 162 | 98 | 38 | 95 |
| GT | 3 | 2 | 2 | 5 |
| TT | 0 | 0 | 0 | 0 |

**Table 2.** (continued)

| | **RESPONDERS** | | **NONRESPONDERS** | |
|---|---|---|---|---|
| | **N** | **%** | **N** | **%** |
| **GENE *S100A9*** | | | | |
| **rs11205276** | | | | |
| GG | 103 | 64 | 25 | 64 |
| GC | 56 | 35 | 14 | 36 |
| CC | 2 | 1 | 0 | 0 |
| **rs3014866*** | | | | |
| TT | 52 | 32 | 14 | 35 |
| TC | 77 | 47 | 22 | 55 |
| CC | 35 | 21 | 4 | 10 |
| **GENE *S100A12*** | | | | |
| **rs724781** | | | | |
| CC | 73 | 46 | 20 | 51 |
| CG | 69 | 43 | 17 | 44 |
| GG | 18 | 11 | 2 | 5 |
| **GENE *S100A8*** | | | | |
| **rs3006488** | | | | |
| AA | 141 | 86 | 33 | 80 |
| AG | 21 | 13 | 7 | 17 |
| GG | 2 | 1 | 1 | 2 |
| **GENE *IL11*** | | | | |
| **rs1126760** | | | | |
| TT | 97 | 60 | 17 | 45 |
| TC | 56 | 35 | 20 | 53 |
| CC | 9 | 6 | 1 | 3 |
| **rs1042506** | | | | |
| TT | 114 | 78 | 24 | 77 |
| TG | 27 | 18 | 7 | 23 |
| GG | 5 | 3 | 0 | 0 |

*CC genotype, p=0.034 OR=3.54 95% CI=1.02-28.82.

**Table 3.** Haplotype frequencies and comparison between Crohn's patients, responders and nonresponders to infliximab therapy, in the chromosomal regions including: A) *TNFAIP6* gene (rs11677200, rs2342910, rs3755480, rs10432475); B) *G0S2* gene (rs4844486, rs1473683); C) *S100A9* and *S100A8* genes (rs11205276, rs3014866, rs724781, and rs3006488) and D) *IL11* gene (rs1126760, rs1042506)

A)

| Haplotyp | Responders | Nonresponders | P value |
|---|---|---|---|
| CAGA | 43.1 | 40.3 | 0.58 |
| TTGA | 18.5 | 18.1 | 0.90 |
| TAGA | 16.6 | 16.8 | 0.99 |
| TAAA | 12.7 | 16.7 | 0.36 |
| TTGG | 8.8 | 8.0 | 0.78 |

B)

| Haplotype | Responders | Nonresponders | P value |
|---|---|---|---|
| CG | 74.1 | 72.8 | 0.76 |
| AG | 25.0 | 24.9 | 0.91 |
| AT | 0.9 | 2.3 | 0.30 |

C)

| Haplotype | Responders | Nonresponders | P value |
|---|---|---|---|
| GCCA | 22.1 | 34.0 | 0.02 |
| GTCA | 24.0 | 19.1 | 0.28 |
| GCGA | 22.7 | 18.8 | 0.42 |
| CTCA | 11.3 | 10.3 | 0.80 |
| CCCA | 4.8 | 4.7 | 1.00 |
| GTGA | 5.5 | 1.7 | 0.39 |
| GCCG | 3.7 | 4.5 | 0.75 |
| GTGG | 1.9 | 3.9 | 0.40 |
| CCGA | 1.5 | 0.9 | 1.00 |
| GTCG | 1.4 | 1.0 | 1.00 |

**Table 3.** (continued)

D)

| Haplotype | Responders | Nonresponders | P value |
|---|---|---|---|
| TT | 77.7 | 72.5 | 0.38 |
| CG | 12.6 | 8.0 | 0.31 |
| CT | 9.6 | 17.8 | 0.06 |
| TG | 0 | 1.6 | 0.18 |

**Table 4.** Haplotype frequencies (haplotypes >2% frequency shown) and comparison between Crohn's patients, responders and nonresponders to infliximab therapy, including all SNPs studied: *IL11* gene (rs1126760, rs1042506), *S100A9* and *S100A8* genes (rs11205276, rs3014866, rs724781, rs3006488), *TNFAIP6* gene (rs11677200, rs2342910, rs3755480, rs10432475) and *G0S2* gene (rs4844486, rs1473683).

| Haplotype | Responders | Nonresponders | P value |
|---|---|---|---|
| TTGCCATTGACG | 8.3 | 7.7 | 0.9822 |
| TTGCGACAGACG | 8.5 | 3.8 | 0.1748 |
| TTGTCATAGACG | 6.6 | 5.9 | 1.0000 |
| TTCTCACAGACG | 5.1 | 5.1 | 1.0000 |
| TTGCCACAGACG | 5.4 | 2.9 | 0.7470 |
| TTGTCACAGACG | 4.5 | 3.3 | 1.0000 |
| TTGCCACAGAAG | 3.8 | 3.7 | 1.0000 |
| TTGCCATAAACG | 1.9 | 6.6 | 0.0532 |
| TTGTCATTGGCG | 2.7 | 1.5 | 1.0000 |
| *CTGCCA*CAGACG | 1.3 | 7.8 | 0.0098 |

**Discussion**

The hierarchical cluster analysis performed by Arijs and collaborators [7] identified a profile with five differentially expressed genes which was claimed to predict response to IFX in CD patients with an overall accuracy of 100%. In that work, the authors validated by qPCR a previously published gene signature which predicted response to IFX in ulcerative colitis patients, [7] but the one discriminating CD patients required corroboration. *IL11* was the only overlapping gene between the two predictive top-five gene sets in both clinical forms of IBD. We aimed at validating the expression signature found for CD through the association study of the reported genes in a totally separated cohort of Spanish CD patients. By doing so, we would ascertain whether any of these genes is causally related with the mechanism of IFX response and might find genetic markers that would be easier to test.

In the independent cohort of CD patients, the tagging approach allowed us to explore a higher genetic variability within the chromosomal regions where the five genes map. One of these five top genes is *TNFAIP6* (tumor necrosis factor, alpha-induced protein 6), which encodes a multifunctional protein with important roles in inflammation and tissue remodeling. It is upregulated in many inflammatory conditions, as rheumatoid arthritis [13] and in colorectal cancer.[14] Another gene previously showing a differential expression profile predictive of response to IFX is *G0S2* (G0/G1 switch 2), involved in lymphocyte cell cycle regulation and found upregulated in rheumatoid arthritis and psoriasis. [15,16] None of the SNPs or haplotypes studied in these two genes evidenced direct association with the response to IFX in Spanish CD patients, probably due to statistical power constraints.

Other two genes included in the predictive expression panel for IFX responsiveness are *S100A8* and *S100A9;* both encode members which belong to the S100 family of calcium-binding proteins and are located in a cluster on chromosome 1q21. Their expression is induced by proinflammatory cytokines such as IL-6 or TNF-□.[17] Calprotectin, the heterodimeric complex of S100A8 and S100A9, shows increased expression as an early step in the neoplastic transformation during colorectal carcinogenesis [18] and it is associated with disease activity in patients with IBD [19] and other inflammatory conditions as rheumatoid arthritis [20] or systemic lupus erythematosus.[21] Moreover, fecal calprotectin concentration is considered a useful surrogate marker for mucosal healing during TNF□-blocking therapy for IBD. [2,22]

In this genetic region, as mentioned, the most abundant haplotype was found associated with response to IFX.

Another gene with reported downregulated expression in IFX responders is *IL11*. This interleukin is a member of the gp130 family of cytokines that stimulates T-cell dependent development of immunoglobulin-producing B cells, and that was once tested as a therapy in CD but appeared to be poor in maintaining remission.[23] In this case, promising association results have been found involving an *IL11* haplotype with decreased frequency in CD responders to IFX.

In our patients, induction therapy with IFX achieved a very good outcome with higher rates of response than those reported in controlled trials.[3,4] The differences between clinical trails and clinical practice are most probably owed to the selection of patients and to the evaluation of response.[24] A high response rate to the three induction doses of IFX was previously reported in a multicenter study performed in our country.[25]

Smoking has been consistently identified as a risk factor for CD. Active smoking in CD patients is associated with worse outcome and increased therapeutic requirements to control the disease.[26] Initial data suggested that active smoking reduced the likeliness to respond to one single IFX induction dose,[27] although larger and more recent studies, including the pivotal controlled studies [28,29] repeatedly found no association between smoking and the response to IFX. Concordantly, in our study with patients who received three induction doses of IFX, the smoking status showed a non-significant difference on response. In a recent study, the combination therapy with infliximab plus azathioprine increased the response rates in CD patients, when compared with IFX monotherapy.[30] The concomitant use of immunomodulators did not influence the rate of response to IFX in our cohort of patients.

Pharmacogenetics has emerged as a promising discipline which opens the possibility of a personalized medicine. However, research has been hampered mainly due to limitations that this kind of studies still shows. Recruitment of a high number of patients with a similar treatment is not an easy task, but difficulty increases because response criteria and clinical data should be centrally monitored. Moreover, the high success rate of some pharmacological therapies as IFX originates a low number of nonresponders, with the consequent decrease in the statistical power to detect differences between responders and nonresponders. The strategy followed in our study

validates the reported gene expression profile predictive of anti-TNF therapy in an independent cohort of Spanish CD patients. Genetic studies stand out as approaches to define pathogenic pathways and ultimately, the integration of genetic together with functional data promotes a clearer understanding of the mechanisms underlying therapeutic pathways. Nevertheless, before the top-five gene signature can be used in clinical practice, its predictive value should be increased by including additional biomarkers of response, but a promising horizon opens regarding classification of response.

**Acknowledgements**

We thank patients and controls for making this study feasible. Angel García and Carmen Martínez provided expert technical assistance. Financial support for the study was provided by: Fondo de Investigaciones Sanitarias FIS (PI08/1636) and Fundación Mutua Madrileña. Elena Urcelay works for the Fundación para la Investigación Biomédica-Hospital Clínico San Carlos (IdISSC).

L.M.M participated in the design of the study, performed the genotyping, analyzed the data, and drafted the manuscript. C.T. participated in the design of the study, and drafted the manuscript. E.U., C.N made substantial contribution to the conception and design of the study, acquisition of data, coordination and drafting of the manuscript. M.B.A., J.L.P., F.B., A.L.S, M.D.M.A, J.P.G., J.L.M., M.D., and E.G.C. participated in the acquisition and interpretation of data. All the authors revised critically the manuscript, gave necessary attention to ensure the integrity of the work presented, and approved the final version.

**References**

1. Abraham C, Cho JH (2009) Inflammatory bowel disease. N Engl J Med 361: 2066-2078.

2. Franke A, McGovern DP, Barrett JC, Wang K, Radford-Smith GL, et al. (2010) Genome-wide meta-analysis increases to 71 the number of confirmed Crohn's disease susceptibility loci. Nat Genet 42: 1118-1125.

3. Rutgeerts P, Feagan BG, Lichtenstein GR, Mayer LF, Schreiber S, et al. (2004) Comparison of scheduled and episodic treatment strategies of infliximab in Crohn's disease. Gastroenterology 126: 402-413.

4. Sands BE, Anderson FH, Bernstein CN, Chey WY, Feagan BG, et al. (2004) Infliximab maintenance therapy for fistulizing Crohn's disease. N Engl J Med 350: 876-885.

5. Feagan BG, Lemann M, Befrits R, Connell W, D'Haens G, et al. (2011) Recommendations for the treatment of Crohn's disease with tumor necrosis factor antagonists: an expert consensus report. Inflamm Bowel Dis 18: 152-160.

6. Cottone M, Criscuoli V (2011) Infliximab to treat Crohn's disease: an update. Clin Exp Gastroenterol 4: 227-238.

7. Arijs I, Quintens R, Van Lommel L, Van Steen K, De Hertogh G, et al. (2010) Predictive value of epithelial gene expression profiles for response to infliximab in Crohn's disease. Inflamm Bowel Dis 16: 2090-2098.

8. Toonen EJ, Gilissen C, Franke B, Kievit W, Eijsbouts AM, et al. (2012) Validation study of existing gene expression signatures for anti-TNF treatment in patients with rheumatoid arthritis. PLoS One 7: e33199.

9. Lennard-Jones JE (1989) Classification of inflammatory bowel disease. Scand J Gastroenterol Suppl 170: 2-6; discussion 16-19.

10. Silverberg MS, Satsangi J, Ahmad T, Arnott ID, Bernstein CN, et al. (2005) Toward an integrated clinical, molecular and serological classification of inflammatory bowel disease: Report of a Working Party of the 2005 Montreal World Congress of Gastroenterology. Can J Gastroenterol 19 Suppl A: 5-36.

11. Harvey RF, Bradshaw JM (1980) A simple index of Crohn's-disease activity. Lancet 1: 514.

12. Colombel JF, Sandborn WJ, Reinisch W, Mantzaris GJ, Kornbluth A, et al. (2010) Infliximab, azathioprine, or combination therapy for Crohn's disease. N Engl J Med 362: 1383-1395.

13. Nagyeri G, Radacs M, Ghassemi-Nejad S, Tryniszewska B, Olasz K, et al. (2011) TSG-6 protein, a negative regulator of inflammatory arthritis, forms a ternary complex with murine mast cell tryptases and heparin. J Biol Chem 286: 23559-23569.
14. Offenberg H, Brunner N, Mansilla F, Orntoft Torben F, Birkenkamp-Demtroder K (2008) TIMP-1 expression in human colorectal cancer is associated with TGF-B1, LOXL2, INHBA1, TNF-AIP6 and TIMP-2 transcript profiles. Mol Oncol 2: 233-240.
15. Koczan D, Guthke R, Thiesen HJ, Ibrahim SM, Kundt G, et al. (2005) Gene expression profiling of peripheral blood mononuclear leukocytes from psoriasis patients identifies new immune regulatory molecules. Eur J Dermatol 15: 251-257.
16. Nakamura N, Shimaoka Y, Tougan T, Onda H, Okuzaki D, et al. (2006) Isolation and expression profiling of genes upregulated in bone marrow-derived mononuclear cells of rheumatoid arthritis patients. DNA Res 13: 169-183.
17. Eggers K, Sikora K, Lorenz M, Taubert T, Moobed M, et al. (2011) RAGE-dependent regulation of calcium-binding proteins S100A8 and S100A9 in human THP-1. Exp Clin Endocrinol Diabetes 119: 353-357.
18. Luley K, Noack F, Lehnert H, Homann N (2011) Local calprotectin production in colorectal cancer and polyps--active neutrophil recruitment in carcinogenesis. Int J Colorectal Dis 26: 603-607.
19. Manolakis AC, Kapsoritakis AN, Tiaka EK, Potamianos SP (2011) Calprotectin, calgranulin C, and other members of the s100 protein family in inflammatory bowel disease. Dig Dis Sci 56: 1601-1611.
20. Hammer HB, Fagerhol MK, Wien TN, Kvien TK (2011) The soluble biomarker calprotectin (a S100 protein) is associated to ultrasonographic synovitis scores and is sensitive to change in patients with rheumatoid arthritis treated with adalimumab. Arthritis Res Ther 13: R178.
21. Lood C, Stenstrom M, Tyden H, Gullstrand B, Kallberg E, et al. (2011) Protein synthesis of the pro-inflammatory S100A8/A9 complex in plasmacytoid dendritic cells and cell surface S100A8/A9 on leukocyte subpopulations in systemic lupus erythematosus. Arthritis Res Ther 13: R60.
22. Molander P, Af Bjorkesten CG, Mustonen H, Haapamaki J, Vauhkonen M, et al. (2012) Fecal calprotectin concentration predicts outcome in inflammatory bowel disease after induction therapy with TNFalpha blocking agents. Inflamm Bowel Dis.
23. Herrlinger KR, Witthoeft T, Raedler A, Bokemeyer B, Krummenerl T, et al. (2006) Randomized, double blind controlled trial of subcutaneous recombinant human

interleukin-11 versus prednisolone in active Crohn's disease. Am J Gastroenterol 101: 793-797.

24. Hanauer SB, Feagan BG, Lichtenstein GR, Mayer LF, Schreiber S, et al. (2002) Maintenance infliximab for Crohn's disease: the ACCENT I randomised trial. Lancet 359: 1541-1549.

25. Gonzalez-Lama Y, Lopez-San Roman A, Marin-Jimenez I, Casis B, Vera I, et al. (2008) Open-label infliximab therapy in Crohn's disease: a long-term multicenter study of efficacy, safety and predictors of response. Gastroenterol Hepatol 31: 421-426.

26. Nos P, Domenech E (2011) Management of Crohn's disease in smokers: is an alternative approach necessary? World J Gastroenterol 17: 3567-3574.

27. Parsi MA, Achkar JP, Richardson S, Katz J, Hammel JP, et al. (2002) Predictors of response to infliximab in patients with Crohn's disease. Gastroenterology 123: 707-713.

28. Fefferman DS, Lodhavia PJ, Alsahli M, Falchuk KR, Peppercorn MA, et al. (2004) Smoking and immunomodulators do not influence the response or duration of response to infliximab in Crohn's disease. Inflamm Bowel Dis 10: 346-351.

29. Orlando A, Colombo E, Kohn A, Biancone L, Rizzello F, et al. (2005) Infliximab in the treatment of Crohn's disease: predictors of response in an Italian multicentric open study. Dig Liver Dis 37: 577-583.

30. Milner CM, Higman VA, Day AJ (2006) TSG-6: a pluripotent inflammatory mediator? Biochem Soc Trans 34: 446-450.

**Discusión general**

En las últimas décadas se han producido grandes avances en la búsqueda de factores genéticos de riesgo a diversas enfermedades autoinmunes, pero es necesario seguir avanzando en este tipo de estudios puesto que todavía existe un elevado porcentaje de riesgo genético sin explicar.

En el caso de la ECe, la región del HLA confiere la mayor susceptibilidad genética a padecer la enfermedad. Aunque las principales variantes causales en esta región se conocen desde hace años, el elevado desequilibrio de ligamiento presente en el HLA hace difícil el estudio de nuevas variantes causales. Los estudios de GWAS han avanzado en la búsqueda de nuevos marcadores centrándose principalmente en las regiones no-HLA, por ello nos propusimos estudiar variantes en la región del HLA que pudieran explicar parte de esa "heredabilidad perdida". Nuestros resultados muestran la presencia de un factor de riesgo adicional en uno de los haplotipos que codifica la molécula DQ2.5 *cis*, el haplotipo ancestral (AH) 8.1. [144]El efecto de este factor sólo se manifiesta en individuos que no presentan una segunda copia del alelo *DQB1*02*. Este factor podría estar también presente en los individuos DQ2.5 *trans* debido a que ambos grupos presentan el mismo riesgo a desarrollar ECe, que es superior al de individuos DQ2.5 codificado por un haplotipo distinto al AH 8.1.

Los loci *HLA-DQA1* y *HLA-DQB1* codifican las cadenas α y β de los heterodímeros de la molécula DQ. La estimulación de las células T por los péptidos derivados del gluten depende tanto del número como del tipo de moléculas DQ expresadas. El factor de riesgo que nosotros detectamos en el AH 8.1 no parece incrementar el riesgo en aquellos individuos que ya presentan riesgo muy alto (DQ2.5*cis* homozigotos o DQ2.5*cis*/ DQ2.2). Las moléculas DQ2.5 se pueden unir con gran afinidad a una amplia variedad de péptidos de gliadinas, mientras que en el caso de DQ2.2 la afinidad es menor. [145] Por lo tanto, un factor adicional en aquellos individuos que poseen dos copias de la molécula DQ2 apenas supone un incremento de riesgo. Mientras que en el caso de aquellos individuos DQ2.5 *cis*, sólo una de las cuatro combinaciones de las cadenas α y β codificaría para DQ2, por lo que un factor de riesgo adicional parece conferir mayor inmunogenicidad. Esto mismo se podría explicar en el contexto de estabilidad en las uniones entre la molécula DQ de las células T con los péptidos de gliadina. Las moléculas DQ2.5 permiten uniones más estables que las DQ2.2 [68], por lo que el factor adicional que detectamos podría

permitir uniones más estables y por tanto presentar mejor los péptidos de gliadina, cosa que también podría estar ocurriendo en aquellos individuos DQ2.5 *trans*.

Estos resultados tienen gran importancia en investigaciones futuras, como son las relacionadas con el bloqueo de moléculas DQ2 [146] y la sensibilización con péptidos de gliadina como tratamiento para la ECe [147], ya que en la aplicación de estos tratamientos debería tenerse en cuenta el perfil genético individual en el HLA y habría que considerar el nuevo factor que nosotros describimos.[148]

Sería interesante llegar a determinar la variante causal presente en el AH8.1. Sin embargo, nuestro tamaño muestral es insuficiente dada la complejidad de la región HLA.

Parte de la "heredabilidad perdida" puede deberse a interacciones epistáticas entre diferentes genes. Nuestro trabajo muestra la interacción de dos polimorfismos en los genes *TLR9* e *IL23R* causante de susceptibilidad a EC, pero no a CU.[149] Estas dos enfermedades tienen perfiles genéticos en parte comunes, pero difieren en la implicación de ciertas rutas. La función de estos dos genes hace que este resultado sea de gran interés. Ambos genes codifican proteínas que juegan un papel importante en el mantenimiento de la homeostasis, por una parte TLR9 participa en el reconocimiento de patógenos extracelulares e IL23R es un receptor involucrado en la ruta de las células Th17, células implicadas en la respuesta a infecciones de patógenos extracelulares, cuya alteración juega un papel importante en el desarrollo de enfermedades autoinmunes. Esta interacción fue descrita inicialmente en población alemana [150], aunque a diferencia de ese estudio, nosotros no observamos interacción entre ninguna de las variantes localizadas en *TLR9* y *NOD2*.

La evidencia de una posible interacción de bacterias intestinales con el hospedador como inicio del desarrollo de la EC, confiere un elevado interés a la interacción entre *IL23R* y *TLR9* en el intento de explicar un posible papel funcional integrador de los dos genes en la patogénesis de la enfermedad. Por una parte TLR9 tiene un papel en el reconocimiento de patógenos y también se ha visto involucrado en la alteración de las células de Paneth, [151] células productoras de péptidos antimicrobianos. Por otra parte, el reconocimiento de patrones bacterianos induce la producción de IL-23 por células del sistema inmunológico innato. Estudios recientes con modelos múridos *tlr9*$^{-}$*/tlr9*$^{-}$ muestran como existe un aumento de los niveles de células Treg y una disminución de las células Th17 y Th1 en esos ratones, sugiriendo que el ADN de las bacterias comensales es un adyuvante natural involucrado en la

inflamación intestinal y en el equilibrio entre células Treg y Th17.[152, 153] Por lo tanto, nuestro trabajo confirma los resultados previamente descritos por Török *et al.*[150] en lo referente a *IL23R* y *TLR9*, pudiendo servir de ayuda clínica, ya que la combinación de ciertos genotipos asociados a una respuesta Th17 podría condicionar la respuesta a futuros tratamientos. En pacientes con EII ya se están utilizando nuevos tratamientos con anticuerpos monoclonales anti-IL23.[122-124]

Nuestros estudios también han ayudado a determinar el papel de las Th17 en la susceptibilidad a padecer la ECe. Trabajos previos revelaron la importancia de las células Th17 en ciertas enfermedades autoinmunes: EII, psoriasis, artritis reumatoide... [154-158] Debido al elevado número de variantes compartidas entre diversas enfermedades de carácter inmunológico quisimos comprobar cuál era el papel de las células Th17 en ECe, ya que estudios previos mostraban un aumento de expresión de ciertas citocinas involucradas en esta ruta en individuos con ECe. [83, 114, 115] Todo ello hacía pensar en el posible papel de las Th17 en la ECe a pesar de no haberse visto asociado previamente en los estudios de GWAS. Para evitar posibles causas de falta de asociación en los GWAS previos, analizamos 101 polimorfismos que cubrían la práctica totalidad de variación genética común presente en 16 genes relacionados con la ruta de las Th17 y además analizamos la posible interacción entre esos genes a la hora de conferir susceptibilidad. Sin embargo, no observamos asociación con ninguno de los SNPs analizados, ni interacciones entre ellos, pudiendo concluir que no existe una implicación genética de las células Th17 en el origen de la enfermedad. [159]

Posteriormente a nuestro estudio, se ha seguido observando la implicación de las células Th17 en ECe. [160, 161] Para conocer la implicación de esta ruta en la patogénesis de la enfermedad analizamos la expresión a nivel intestinal en pacientes con ECe. Además, quisimos ver esto mismo en CU, puesto que aunque el número de estudios no es tan elevado como en el caso de EC, está clara su implicación en el origen y patogénesis de la enfermedad. Observamos una diferente expresión de los genes *IL23A* y *JAK2* entre pacientes con ECe y controles. Además obtuvimos resultados interesantes en los pacientes con ECe según la edad de debut: observamos una mayor expresión de *IL6* y *RHBDD3* en pacientes adultos que en niños, en estos últimos la expresión era similar a la de controles. No existen estudios previos que hayan realizado este tipo de comparación dependiendo de la edad de diagnóstico. Nuestros resultados indican que la respuesta inmune que desencadena la enfermedad puede presentar diferencias entre aquellas personas que debutan a edad temprana o

adulta. Por lo tanto, es muy interesante aumentar el número de pacientes en nuestro estudio y llegar a concluir un perfil de expresión diferencial entre adultos y niños, que podría estar relacionado con las diferencias en la presentación clínica entre ambos grupos.[162]

Entre nuestras muestras pareadas de tejido inflamado y sano en CU, observamos una mayor expresión de *IL6, CCR6* y *SMAD3* en tejido inflamado. Nuestros resultados también muestran variación en otros genes, aunque no llega a ser significativa (*IL23R*, *IL21* y *JAK2*). Un estudio reciente de metanálisis en EII muestra niveles elevados de *IL6* y *CCR6* en tejido inflamado de pacientes con CU, confirmándose así esos resultados.[163] En el caso del gen *SMAD3* son necesarios estudios adicionales que confirmen nuestro resultado. Nuestro trabajo apoya la idea de la importancia de las Th17 en la CU y como existe un aumento de expresión de ciertos genes relacionados con su diferenciación o proliferación. La asociación de los niveles de expresión con inflamación puede ser consecuencia de la propia inflamación o causa de la enfermedad. Para aclarar este punto se necesitan estudios comparativos con muestras de colon en individuos sanos que determinen cuáles son los niveles basales de expresión en población general y ver si existen diferencias con tejido inflamado. En aquellos genes en los que se concluya implicación en el origen de la enfermedad, sería interesante conocer las variantes genéticas específicas que contribuyen a alterar la expresión.

Los estudios de expresión en tejido inflamado pueden ser interesantes para comprender mejor la enfermedad, incluso llegar a servir de biomarcadores para realizar un diagnóstico diferencial entre enfermedades, por ejemplo entre CU y EC. Grandlund *et al*. muestran como *IL23A* se expresa en mayor medida en CU que EC, mientras que el resto de las 17 citocinas estudiadas aumentan su expresión en ambas enfermedades por igual. [163] Grandlund *et al*. sugieren por tanto que la expresión de *IL23A* podría servir en el diagnóstico diferencial entre las dos formas de EII.

La diferencia en el perfil de expresión en ECe y CU puede deberse a una implicación diferente de las células Th17 en ambas enfermedades. La EII se ha asociado a diversos genes relacionados con la ruta de las Th17, mientras que en el caso de ECe se ha observado apenas asociación. Estudios recientes muestran que las células Th17 son células inestables que pueden transformarse en otro subtipo de célula Th, su habilidad para convertirse y producir un perfil de citocinas distinto parece que ayuda en el balance homeostático y que depende de variaciones ambientales.[164,

165] Esa plasticidad dificulta su estudio y puede que las diferencias entre ambas enfermedades se deban a esa propia plasticidad de las células. Algunos estudios muestran como existe un diferente patrón tanto de activación como de actuación de las células Th17 en ECe, [115] aunque son necesarios más estudios que expliquen esa plasticidad.

La gran variabilidad genética entre individuos con una misma enfermedad hizo pensar que la respuesta a las diferentes terapias podía verse afectada por la genética del paciente. De esta manera surgieron los estudios farmacogenéticos, que pretenden servir para predecir la respuesta al tratamiento.

En EC, cuando los fármacos inmunomoduladores no son eficaces, se administra IFX, un tratamiento biológico de gran eficacia en pacientes refractarios y con fístulas. En este trabajo nos planteamos avanzar en los estudios de farmacogenética en individuos con EC tratados con IFX. De esta manera nuestro objetivo fue buscar ciertos marcadores genéticos que ayuden a predecir la respuesta al tratamiento. Para ello estudiamos los genes *TNFRSF1B* y *TNFRSF1A* y observamos la asociación de un haplotipo en *TNFRSF1B* cuya frecuencia era mayor en no respondedores frente a respondedores. Este tipo de asociación se vio previamente en población japonesa [166] y con nuestros resultados se confirma la implicación de *TNFRSF1B* en la respuesta al tratamiento con IFX.[167] *TNFRSF1B* es un receptor necesario en la función biológica del TNF, su unión interviene en la activación de rutas proinflamatorias y en la programación de muerte celular. En pacientes de EC con afectación colónica y que no responden a IFX se han observado niveles elevados de *TNFRSF1B*.[168] Es necesario continuar estos estudios aumentando el número de pacientes para ver cuál es la verdadera variante causal implicada en *TNFRSF1B* y así entender cuál es el mecanismo implicado en la respuesta a IFX.

Los niveles de expresión de ciertos genes involucrados en la vía de acción del fármaco o en la patogénesis de la enfermedad pueden variar entre individuos que no responden frente a los respondedores y así definir específicamente el grupo de individuos a los que dirigir el tratamiento.

Los niveles de expresión de cinco genes: *TNFAIP6*, *IL11*, *G0S2*, *S100A8* y *S100A9*, parecen predecir la respuesta a IFX, discriminando hasta el 100% de pacientes con afectación colónica. [168] Nuestros resultados muestran como una combinación alélica de estos cinco genes puede ser válida para diferenciar la respuesta a IFX, puesto que aparece en casi un 8% de los no respondedores frente al 1% de los

respondedores.[169] Nuestros resultados confirman que los cambios de expresión observado por Arijs *et al*. no son consecuencia del tratamiento con IFX, sino de la propia genética del individuo que puede considerarse como marcador predictivo de la respuesta al tratamiento. Es importante que estos resultados de nuestros estudios se repliquen en una muestra independiente para ratificar su validez.

Es necesario definir marcadores predictivos de respuesta a IFX para elegir una terapia adecuada para cada paciente maximizando la eficacia y minimizando la posible toxicidad al fármaco. El tratamiento precoz incidirá en una mejor calidad de vida del paciente con la repercusión económica adicional, ajustando el gasto que en estos tratamientos es importante por su elevado coste. El estudio de marcadores que puedan anticipar el grado de respuesta a ciertos tratamientos sigue siendo necesario para llegar a aplicar en la práctica clínica un tratamiento individualizado.

En conclusión, nuestro trabajo avanza en el conocimiento de las bases genéticas de ECe y EII mediante estudios basados en análisis estratificados como es el desarrollo en la región HLA en ECe y en interacciones epistáticas, análisis que no están incluidos en los estudios de GWAS. También avanzamos en la búsqueda de marcadores farmacogenéticos para definir una respuesta en pacientes con EC tratados con IFX; estudiando un grupo amplio y homogéneo de pacientes unificados por criterios de respuesta al tratamiento.

**Situación actual y perpectivas futuras**

Los diferentes avances en la búsqueda de factores asociados a las enfermedades autoinmunes han puesto de manifiesto nuevas rutas que han ayudado a conocer mejor la patogénesis de estas enfermedades. Además algunas de estas rutas pueden tener relevancia clínica, puesto que abren la posibilidad de encontrar nuevos biomarcadores tanto para el diagnóstico de la EII como de la ECe. Y lo más importante es que las diferentes variantes causales u otras moléculas identificas pueden representar nuevas posibles dianas terapéuticas, no sólo para la enfermedad en que se ha encontrado, en este caso EII o ECe, sino para otras enfermedades autoinmunes relacionadas. [74]

Al principio se pensó que la clave de esa falta de "heredabilidad perdida" podía estar en *copy number varitation* (CNV)[170] realizándose diversos estudios sin obtener grandes resultados. Actualmente se están desarrollando nuevas tecnologías de secuenciación masiva para identificar variantes muy poco frecuentes que no están

incluidas en los estudios previos de GWAS [171] y tratar de explicar la "heredabilidad perdida". La secuenciación de genoma (WGS) y del exoma (WES) son dos de las principales estrategias a seguir en un futuro. En enfermedades mendelianas han sido de gran utilidad para identificar nuevas variables.[172] En el caso de WES, está diseñada para identificar variantes localizadas en los codones de exones y que puedan afectar a la secuencia de la proteína (variantes no sinónimas), es decir, se centra en regiones codificantes que puedan alterar la función de la proteína. Mientras que en el caso de WGS nos permite detectar variantes en exones, intrones y regiones intergénicas que puedan estar regulando además la expresión génica. El principal inconveniente de estas tecnologías es su precio. Sin embargo, estudios preliminares de la secuenciación del exoma en 25 regiones previamente asociadas a enfermedades complejas mediadas por el sistema inmunológico han revelado sólo una variante adicional[63] Por otro lado, hay que tener especial cuidado en los estudios de secuenciación, ya que las variantes raras pueden ser específicas de una población determinada. Rivas *et al*. en 2011 con estudios de secuenciación en EC identificaron nuevas variantes independientes asociadas a la enfermedad en *NOD2, CARD9, IL23, IL18RAP, CUL2, C1orf106, PTPN22 y MUC19*, que podrían ayudar a priorizar estudios funcionales y modelos animales predictivos. [173] El último trabajo de secuenciación en CU ha identificado tres nuevas variantes en *CARD9*, *IL23R* y *RNF186* como posible variantes causales en CU, relacionando la coexpresión de *RNF186* en el intestino delgado y colon con el gen previamente asociado a CU *HNF4A*. [75]

También es posible aumentar la resolución de ciertos estudios con la imputación de los *arrays* de GWAS usando como referencia el proyecto de los 1000 genomas, [174]también éste tipo de herramienta se puede aplicar en el Proyecto Inmunochip. Estos estudios son de gran utilidad para la búsqueda de nuevas asociaciones. Un estudio reciente llevado a cabo por el *Wellcome Trust Case-control Consortium* llega a determinar dos nuevos loci: *IL2RA* en DM1 y *CDKN2B* en DM2, y variantes potencialmente causales de dos loci previamente conocidos en DM1 y EC que pueden ayudar a definir la biología de la enfermedad. [175]

Otra de las maneras de priorizar genes asociados en GWAS es incorporar datos funcionales, desde datos de expresión hasta datos de interacciones proteína-proteína. Ricano-Ponce *et al*. con las variantes asociadas previamente a las 12 enfermedades inmuno-relacionadas incluidas en el Inmunochip observaron como un 3% de las

variantes asociadas se localizaban en exones, 42% en intrones, y el resto (55%) en regiones intergénicas y no codificantes. Usando los resultados del ENCODE (*Encyclopedia of DNA Elements*) es posible posicionar los SNPs en las regiones reguladoras del genoma. Riaño-Ponce *et al.* han demostrado que la mayoría de las variantes asociadas son potencialmente reguladoras, determinando que afectan a la expresión de genes cercanos. [176] Esta última idea hace que esté cobrando más importancia el estudio de moléculas reguladoras como son *small non-coding RNA* (microRNA, miRNA) y *long intergenic non-coding RNA* (lincRNA) en las enfermedades autoinmunes. Se trata de moléculas de RNA reguladoras del genoma que no se transcriben. La función de miRNA y lincRNA puede alterarse debido a variaciones genéticas y afectar a la transcripción de genes involucrados en la enfermedad. [177, 178]

Otra manera de abordar la búsqueda de variantes genéticas de riesgo es mediante la realización de meta-análisis, en los que se incluyen datos de diversos estudios y que pueden incluir distintas poblaciones o enfermedades relacionadas. Es una buena manera de aumentar la potencia estadística para buscar ese porcentaje de heredabilidad aún por explicar y poder identificar nuevas variantes compartidas con un efecto bajo. Un ejemplo de meta-análisis es el llevado a cabo por Festen *et al*. en EC y ECe, en el que identificaron dos nuevos loci compartidos: *PUS10* y *TAGAP*. [179] Posteriormente *TAGAP* se asoció también a EM. [180]

Los resultados de este tipo de análisis son de gran utilidad para la búsqueda de tratamientos comunes a diversas enfermedades autoinmunes. Aunque también lo son aquellos genes específicos de cada enfermedad, ya que estarán involucrados en una ruta de señalización única, siendo de gran interés descifrar los mecanismos en los que se implique en la enfermedad para una posible prevención y búsqueda de tratamientos específicos.

Por otra parte, el microbioma parece jugar un papel importante en las enfermedades intestinales autoinmunes al mediar en la respuesta inmune frente a antígenos inocuos en el medio intestinal. En la EII ya se postula que la flora intestinal tenga un papel etiológico en el desarrollo de la enfermedad. En la EII diversos factores genéticos están relacionados con el cambio en la microbiota comensal.[181] También se ha observado en los diferentes grupos de pacientes con ECe una diferencia en los componentes de la microbiota: con manifestaciones extraintestinales, en dieta libre de gluten o en ECe activa. [182-184] Es necesario continuar con los estudios en este

campo, puesto que presenta una gran complejidad por sus relaciones con factores ambientales y con variaciones en el genoma del individuo. Un objetivo fundamental en el estudio de enfermedades multifactoriales es entender cómo las diferentes rutas funcionales conocidas a través de estudios genéticos se integran con el conocimiento de microorganismos o estímulos ambientales que puedan estar implicados en estas enfermedades.

El objetivo de los estudios genéticos no es sólo describir variantes de riesgo presentes en el genoma, sino también entender los mecanismos biológicos que contribuyen al desarrollo de la enfermedad. Las nuevas aproximaciones parecen prometedoras para avanzar en este objetivo pero no hay que olvidar que estudios bien diseñados con pacientes bien caracterizados como son los desarrollados en este trabajo pueden contribuir al mejor conocimiento de las enfermedades complejas.

Este trabajo define nuevos factores de susceptibilidad implicados tanto en EII como en ECe y puede ayudar a comprender los mecanismos moleculares involucrados en estas enfermedades. Además, define marcadores genéticos que predicen una pérdida de respuesta a IFX, uno de los tratamientos de elección en la EC. Las conclusiones de este trabajo son:

1. Existe un factor genético de susceptibilidad a ECe localizado en el HLA y presente en el haplotipo ancestral 8.1. El riesgo conferido por este factor sólo se observa cuando los individuos no portan un segundo alelo *DQB1*02*. Este factor de riesgo podría también estar presente en individuos DQ2.5 *trans*, dado el riesgo similar a padecer ECe que presentan ambos conjuntos de individuos.

2. La interacción entre variantes genéticas presentes en *TLR9* e *IL23R* contribuye al desarrollo de la enfermedad de Crohn en población blanca. Por el contrario, la interacción previamente descrita en población alemana entre *IL23R* y *NOD2* no parece influir.

3. Los principales genes implicados en la respuesta inmune tipo Th17 no parecen estar asociados al riesgo a padecer enfermedad celiaca en población española. Por el contrario, la expresión alterada de los genes *IL23A*, *JAK2*, *IL6* y *RHBDD3* en pacientes de ECe sugiere que la ruta Th17 está implicada en la patogenia de la enfermedad.

4. Dependiendo de la edad de debut de la enfermedad celiaca puede existir un perfil de expresión génica diferente, como sugieren nuestros resultados en *IL6* y *RHBDD3*.

5. La expresión de *CCR6*, *IL6* y *SMAD3*, todos ellos implicados en la respuesta tipo Th17, aumenta en regiones inflamadas de pacientes con colitis ulcerosa activa.

6. La ruta Th17 parece importante en la patogenia de la enfermedad celiaca y colitis ulcerosa como se refleja por la expresión alterada de ciertos genes que, sin embargo, parece diferir entre ambas enfermedades.

7. El gen *TNFRSF1B* está implicado en la respuesta a infliximab en individuos con enfermedad de Crohn.

8. La expresión diferencial de los genes *TNFAIP6*, *S100A8*, *IL11*, *G0S2* y *S100A* observada entre pacientes de enfermedad de Crohn respondedores y no respondedores a infliximab parece venir determinada por variantes específicas presentes en esos genes.

*Conclusions*

This work associates new risk factors in IBD and CeD and it can help to understand the molecular mechanisms involved in the disease. In addition, we found genetic markers which predict loss of response to IFX, one of the elections to treat CD patients. The conclusions of this work are:

1. There is a risk variant for celiac disease in the HLA region which is present in the ancestral haplotype 8.1. The risk conferred by this factor is only observed in individuals who do not carry a second allele *DQB1*02*. This factor can be also present in DQ2.5 *trans* individuals, due to the similar CeD risk in both groups.

2. The interaction between genetic variants in *TLR9* and *IL23R* plays a role in Crohn´s disease susceptibility in white population. By the contrary, the previously described interaction in German population between *IL23R* and *NOD2* does not seem to influence disease risk.

3. The main genes involved in the Th17 pathway do not seem to influence celiac disease risk in Spanish population. By the contrary, our results show the altered expression of *IL23A*, *JAK2*, *IL6* and *RHBDD3* in celiac disease patients, suggesting that the Th17 pathway is involved in CeD pathogeny.

4. Depending on the age of onset in celiac disease a different expression profile may exist, as our results in *IL6* and *RHBDD3* genes suggest.

5. The expression levels of the *CCR6*, *IL6* and *SMAD3* genes, which are involved in the Th17 immune response, are elevated in inflamed regions of ulcerative colitis patients.

6. The Th17 pathway can be important in the pathogeny of celiac disease and ulcerative colitis as a result of the altered expression observed in several genes, which seem to differ between both diseases.

7. The *TNFRSF1B* gene is involved in the response to infliximab of Crohn's disease patients.

8. The observed differential expression of the *TNFAIP6*, *S100A8*, *IL11*, *G0S2* y *S100A* genes between Crohn´s disease responders and non-responders to infliximab seems to be a consequence of specific genetic variants in those genes.

1. Abraham C, Cho JH. Inflammatory bowel disease. N Engl J Med 361(21):2066-78, 2009.
2. Danese S, Fiocchi C. Ulcerative colitis. N Engl J Med 365(18):1713-25, 2011.
3. Loftus EV, Jr. Clinical epidemiology of inflammatory bowel disease: Incidence, prevalence, and environmental influences. Gastroenterology 126(6):1504-17, 2004.
4. Logan I, Bowlus CL. The geoepidemiology of autoimmune intestinal diseases. Autoimmun Rev 9(5):A372-8, 2010.
5. Loftus EV, Jr., Sandborn WJ. Epidemiology of inflammatory bowel disease. Gastroenterol Clin North Am 31(1):1-20, 2002.
6. Bardhan KD, Simmonds N, Royston C, Dhar A, Edwards CM. A United Kingdom inflammatory bowel disease database: making the effort worthwhile. J Crohns Colitis 4(4):405-12, 2010.
7. Bernstein CN, Blanchard JF, Rawsthorne P, Yu N. The prevalence of extraintestinal diseases in inflammatory bowel disease: a population-based study. Am J Gastroenterol 96(4):1116-22, 2001.
8. Silverberg MS, Satsangi J, Ahmad T, Arnott ID, Bernstein CN, Brant SR, et al. Toward an integrated clinical, molecular and serological classification of inflammatory bowel disease: Report of a Working Party of the 2005 Montreal World Congress of Gastroenterology. Can J Gastroenterol 19 Suppl A:5-36, 2005.
9. Catassi C, Anderson RP, Hill ID, Koletzko S, Lionetti E, Mouane N, et al. World perspective on celiac disease. J Pediatr Gastroenterol Nutr 55(5):494-9, 2012.
10. Fasano A, Berti I, Gerarduzzi T, Not T, Colletti RB, Drago S, et al. Prevalence of celiac disease in at-risk and not-at-risk groups in the United States: a large multicenter study. Arch Intern Med 163(3):286-92, 2003.

11. Mustalahti K, Catassi C, Reunanen A, Fabiani E, Heier M, McMillan S, et al. The prevalence of celiac disease in Europe: results of a centralized, international mass screening project. Ann Med 42(8):587-95, 2010.

12. Hin H, Bird G, Fisher P, Mahy N, Jewell D. Coeliac disease in primary care: case finding study. Bmj 318(7177):164-7, 1999.

13. Marine M, Farre C, Alsina M, Vilar P, Cortijo M, Salas A, et al. The prevalence of coeliac disease is significantly higher in children compared with adults. Aliment Pharmacol Ther 33(4):477-86, 2011.

14. Salardi S, Volta U, Zucchini S, Fiorini E, Maltoni G, Vaira B, et al. Prevalence of celiac disease in children with type 1 diabetes mellitus increased in the mid-1990 s: an 18-year longitudinal study based on anti-endomysial antibodies. J Pediatr Gastroenterol Nutr 46(5):612-4, 2008.

15. Valentino R, Savastano S, Tommaselli AP, Dorato M, Scarpitta MT, Gigante M, et al. Prevalence of coeliac disease in patients with thyroid autoimmunity. Horm Res 51(3):124-7, 1999.

16. Korponay-Szabo IR, Dahlbom I, Laurila K, Koskinen S, Woolley N, Partanen J, et al. Elevation of IgG antibodies against tissue transglutaminase as a diagnostic tool for coeliac disease in selective IgA deficiency. Gut 52(11):1567-71, 2003.

17. Kagnoff MF. Celiac disease: pathogenesis of a model immunogenetic disease. J Clin Invest 117(1):41-9, 2007.

18. Catassi C, Ratsch IM, Fabiani E, Rossini M, Bordicchia F, Candela F, et al. Coeliac disease in the year 2000: exploring the iceberg. Lancet 343(8891):200-3, 1994.

19. Ludvigsson JF, Leffler DA, Bai JC, Biagi F, Fasano A, Green PH, et al. The Oslo definitions for coeliac disease and related terms. Gut 62(1):43-52, 2013.

20. Leone V, Chang EB, Devkota S. Diet, microbes, and host genetics: the perfect storm in inflammatory bowel diseases. J Gastroenterol 48(3):315-21, 2013.

21. Berry D, Reinisch W. Intestinal microbiota: A source of novel biomarkers in inflammatory bowel diseases? Best Pract Res Clin Gastroenterol 27(1):47-58, 2013.

22. Gent AE, Hellier MD, Grace RH, Swarbrick ET, Coggon D. Inflammatory bowel disease and domestic hygiene in infancy. Lancet 343(8900):766-7, 1994.

23. Falkow S. Molecular Koch's postulates applied to microbial pathogenicity. Rev Infect Dis 10 Suppl 2:S274-6, 1988.

24. Tamboli CP, Neut C, Desreumaux P, Colombel JF. Dysbiosis as a prerequisite for IBD. Gut 53(7):1057, 2004.

25. Glasser AL, Boudeau J, Barnich N, Perruchot MH, Colombel JF, Darfeuille-Michaud A. Adherent invasive Escherichia coli strains from patients with Crohn's disease survive and replicate within macrophages without inducing host cell death. Infect Immun 69(9):5529-37, 2001.

26. Rook GA, Brunet LR. Microbes, immunoregulation, and the gut. Gut 54(3):317-20, 2005.

27. Lakatos PL, Szamosi T, Lakatos L. Smoking in inflammatory bowel diseases: good, bad or ugly? World J Gastroenterol 13(46):6134-9, 2007.

28. Cosnes J. Tobacco and IBD: relevance in the understanding of disease mechanisms and clinical practice. Best Pract Res Clin Gastroenterol 18(3):481-96, 2004.

29. Koutroubakis IE, Vlachonikolis IG. Appendectomy and the development of ulcerative colitis: results of a metaanalysis of published case-control studies. Am J Gastroenterol 95(1):171-6, 2000.

30. Ananthakrishnan AN. Environmental risk factors for inflammatory bowel disease. Gastroenterol Hepatol (N Y) 9(6):367-74, 2013.

31. Stene LC, Honeyman MC, Hoffenberg EJ, Haas JE, Sokol RJ, Emery L, et al. Rotavirus infection frequency and risk of celiac disease autoimmunity in early childhood: a longitudinal study. Am J Gastroenterol 101(10):2333-40, 2006.

32. Malamut G, Afchain P, Verkarre V, Lecomte T, Amiot A, Damotte D, et al. Presentation and long-term follow-up of refractory celiac disease: comparison of type I with type II. Gastroenterology 136(1):81-90, 2009.

33. Helbig I, Hodge SE, Ottman R. Familial cosegregation of rare genetic variants with disease in complex disorders. Eur J Hum Genet 21(4):444-50, 2013.

34. Cooney R, Jewell D. The genetic basis of inflammatory bowel disease. Dig Dis 27(4):428-42, 2009.

35. Gaya DR, Russell RK, Nimmo ER, Satsangi J. New genes in inflammatory bowel disease: lessons for complex diseases? Lancet 367(9518):1271-84, 2006.

36. Forabosco P, Neuhausen SL, Greco L, Naluai AT, Wijmenga C, Saavalainen P, et al. Meta-analysis of genome-wide linkage studies in celiac disease. Hum Hered 68(4):223-30, 2009.

37. Wolters VM, Wijmenga C. Genetic background of celiac disease and its clinical implications. Am J Gastroenterol 103(1):190-5, 2008.

38. Altmuller J, Palmer LJ, Fischer G, Scherb H, Wjst M. Genomewide scans of complex human diseases: true linkage is hard to find. Am J Hum Genet 69(5):936-50, 2001.

39. Consortium WTCC. Genome-wide association study of 14,000 cases of seven common diseases and 3,000 shared controls. Nature 447(7145):661-78, 2007.

40. Xavier RJ, Rioux JD. Genome-wide association studies: a new window into immune-mediated diseases. Nat Rev Immunol 8(8):631-43, 2008.

41. Spehlmann ME, Begun AZ, Burghardt J, Lepage P, Raedler A, Schreiber S. Epidemiology of inflammatory bowel disease in a German twin cohort: results of a nationwide study. Inflamm Bowel Dis 14(7):968-76, 2008.

42. Hugot JP, Chamaillard M, Zouali H, Lesage S, Cezard JP, Belaiche J, et al. Association of NOD2 leucine-rich repeat variants with susceptibility to Crohn's disease. Nature 411(6837):599-603, 2001.

43. Ogura Y, Bonen DK, Inohara N, Nicolae DL, Chen FF, Ramos R, et al. A frameshift mutation in NOD2 associated with susceptibility to Crohn's disease. Nature 411(6837):603-6, 2001.

44. van der Linde K, Kuipers EJ, de Rooij FW, Wilson JH. [From gene to disease; 'frame shift'-mutation in the CARD15-gene and Crohn's disease]. Ned Tijdschr Geneeskd 146(52):2539-42, 2002.

45. Mendoza JL, Murillo LS, Fernandez L, Pena AS, Lana R, Urcelay E, et al. Prevalence of mutations of the NOD2/CARD15 gene and relation to phenotype in Spanish patients with Crohn disease. Scand J Gastroenterol 38(12):1235-40, 2003.

46. Bouma G, Strober W. The immunological and genetic basis of inflammatory bowel disease. Nat Rev Immunol 3(7):521-33, 2003.

47. Cuthbert AP, Fisher SA, Mirza MM, King K, Hampe J, Croucher PJ, et al. The contribution of NOD2 gene mutations to the risk and site of disease in inflammatory bowel disease. Gastroenterology 122(4):867-74, 2002.

48. King K, Sheikh MF, Cuthbert AP, Fisher SA, Onnie CM, Mirza MM, et al. Mutation, selection, and evolution of the Crohn disease susceptibility gene CARD15. Hum Mutat 27(1):44-54, 2006.

49. Satsangi J, Welsh KI, Bunce M, Julier C, Farrant JM, Bell JI, et al. Contribution of genes of the major histocompatibility complex to

susceptibility and disease phenotype in inflammatory bowel disease. Lancet 347(9010):1212-7, 1996.

50. Roussomoustakaki M, Satsangi J, Welsh K, Louis E, Fanning G, Targan S, et al. Genetic markers may predict disease behavior in patients with ulcerative colitis. Gastroenterology 112(6):1845-53, 1997.

51. Silverberg MS, Mirea L, Bull SB, Murphy JE, Steinhart AH, Greenberg GR, et al. A population- and family-based study of Canadian families reveals association of HLA DRB1*0103 with colonic involvement in inflammatory bowel disease. Inflamm Bowel Dis 9(1):1-9, 2003.

52. Fernandez L, Mendoza JL, Martinez A, Urcelay E, Fernandez-Arquero M, Garcia-Paredes J, et al. IBD1 and IBD3 determine location of Crohn's disease in the Spanish population. Inflamm Bowel Dis 10(6):715-22, 2004.

53. Fisher SA, Tremelling M, Anderson CA, Gwilliam R, Bumpstead S, Prescott NJ, et al. Genetic determinants of ulcerative colitis include the ECM1 locus and five loci implicated in Crohn's disease. Nat Genet 40(6):710-2, 2008.

54. Anderson CA, Boucher G, Lees CW, Franke A, D'Amato M, Taylor KD, et al. Meta-analysis identifies 29 additional ulcerative colitis risk loci, increasing the number of confirmed associations to 47. Nat Genet 43(3):246-52, 2011.

55. Franke A, McGovern DP, Barrett JC, Wang K, Radford-Smith GL, Ahmad T, et al. Genome-wide meta-analysis increases to 71 the number of confirmed Crohn's disease susceptibility loci. Nat Genet 42(12):1118-25, 2010.

56. Jostins L, Ripke S, Weersma RK, Duerr RH, McGovern DP, Hui KY, et al. Host-microbe interactions have shaped the genetic architecture of inflammatory bowel disease. Nature 491(7422):119-24, 2012.

57. Julia A, Domenech E, Ricart E, Tortosa R, Garcia-Sanchez V, Gisbert JP, et al. A genome-wide association study on a southern European population identifies a new Crohn's disease susceptibility locus at RBX1-EP300. Gut 62(10):1440-5, 2013.

58. Greco L, Romino R, Coto I, Di Cosmo N, Percopo S, Maglio M, et al. The first large population based twin study of coeliac disease. Gut 50(5):624-8, 2002.

59. van Heel DA, Franke L, Hunt KA, Gwilliam R, Zhernakova A, Inouye M, et al. A genome-wide association study for celiac disease identifies risk variants in the region harboring IL2 and IL21. Nat Genet 39(7):827-9, 2007.

60. Hunt KA, Zhernakova A, Turner G, Heap GA, Franke L, Bruinenberg M, et al. Newly identified genetic risk variants for celiac disease related to the immune response. Nat Genet 40(4):395-402, 2008.

61. Dubois PC, Trynka G, Franke L, Hunt KA, Romanos J, Curtotti A, et al. Multiple common variants for celiac disease influencing immune gene expression. Nat Genet 42(4):295-302, 2010.

62. Trynka G, Hunt KA, Bockett NA, Romanos J, Mistry V, Szperl A, et al. Dense genotyping identifies and localizes multiple common and rare variant association signals in celiac disease. Nat Genet 43(12):1193-201, 2011.

63. Hunt KA, Mistry V, Bockett NA, Ahmad T, Ban M, Barker JN, et al. Negligible impact of rare autoimmune-locus coding-region variants on missing heritability. Nature 498(7453):232-5, 2013.

64. Sollid LM, Markussen G, Ek J, Gjerde H, Vartdal F, Thorsby E. Evidence for a primary association of celiac disease to a particular HLA-DQ alpha/beta heterodimer. J Exp Med 169(1):345-50, 1989.

65. van Belzen MJ, Koeleman BP, Crusius JB, Meijer JW, Bardoel AF, Pearson PL, et al. Defining the contribution of the HLA region to

cis DQ2-positive coeliac disease patients. Genes Immun 5(3):215-20, 2004.

66. Louka AS, Nilsson S, Olsson M, Talseth B, Lie BA, Ek J, et al. HLA in coeliac disease families: a novel test of risk modification by the 'other' haplotype when at least one DQA1*05-DQB1*02 haplotype is carried. Tissue Antigens 60(2):147-54, 2002.

67. Karell K, Louka AS, Moodie SJ, Ascher H, Clot F, Greco L, et al. HLA types in celiac disease patients not carrying the DQA1*05-DQB1*02 (DQ2) heterodimer: results from the European Genetics Cluster on Celiac Disease. Hum Immunol 64(4):469-77, 2003.

68. Fallang LE, Bergseng E, Hotta K, Berg-Larsen A, Kim CY, Sollid LM. Differences in the risk of celiac disease associated with HLA-DQ2.5 or HLA-DQ2.2 are related to sustained gluten antigen presentation. Nat Immunol 10(10):1096-101, 2009.

69. Henderson KN, Tye-Din JA, Reid HH, Chen Z, Borg NA, Beissbarth T, et al. A structural and immunological basis for the role of human leukocyte antigen DQ8 in celiac disease. Immunity 27(1):23-34, 2007.

70. Tjon JM, van Bergen J, Koning F. Celiac disease: how complicated can it get? Immunogenetics 62(10):641-51, 2010.

71. Husby S, Koletzko S, Korponay-Szabo IR, Mearin ML, Phillips A, Shamir R, et al. European Society for Pediatric Gastroenterology, Hepatology, and Nutrition guidelines for the diagnosis of coeliac disease. J Pediatr Gastroenterol Nutr 54(1):136-60, 2012.

72. Baschal EE, Aly TA, Jasinski JM, Steck AK, Noble JA, Erlich HA, et al. Defining multiple common "completely" conserved major histocompatibility complex SNP haplotypes. Clin Immunol 132(2):203-14, 2009.

73. de la Concha EG, Cavanillas ML, Cenit MC, Urcelay E, Arroyo R, Fernandez O, et al. DRB1*03:01 haplotypes: differential

contribution to multiple sclerosis risk and specific association with the presence of intrathecal IgM bands. PLoS One 7(2):e31018, 2012.

74. Kumar V, Wijmenga C, Withoff S. From genome-wide association studies to disease mechanisms: celiac disease as a model for autoimmune diseases. Semin Immunopathol 34(4):567-80, 2012.

75. Beaudoin M, Goyette P, Boucher G, Lo KS, Rivas MA, Stevens C, et al. Deep Resequencing of GWAS Loci Identifies Rare Variants in CARD9, IL23R and RNF186 That Are Associated with Ulcerative Colitis. PLoS Genet 9(9), 2013.

76. Cordell HJ. Epistasis: what it means, what it doesn't mean, and statistical methods to detect it in humans. Hum Mol Genet 11(20):2463-8, 2002.

77. Zuk O, Hechter E, Sunyaev SR, Lander ES. The mystery of missing heritability: Genetic interactions create phantom heritability. Proc Natl Acad Sci U S A 109(4):1193-8, 2012.

78. Abadie V, Sollid LM, Barreiro LB, Jabri B. Integration of genetic and immunological insights into a model of celiac disease pathogenesis. Annu Rev Immunol 29:493-525, 2011.

79. Evans DM, Spencer CC, Pointon JJ, Su Z, Harvey D, Kochan G, et al. Interaction between ERAP1 and HLA-B27 in ankylosing spondylitis implicates peptide handling in the mechanism for HLA-B27 in disease susceptibility. Nat Genet 43(8):761-7, 2011.

80. Strange A, Capon F, Spencer CC, Knight J, Weale ME, Allen MH, et al. A genome-wide association study identifies new psoriasis susceptibility loci and an interaction between HLA-C and ERAP1. Nat Genet 42(11):985-90, 2010.

81. Handel AE, Ebers GC, Ramagopalan SV. Epigenetics: molecular mechanisms and implications for disease. Trends Mol Med 16(1):7-16, 2010.

82. Trynka G, Wijmenga C, van Heel DA. A genetic perspective on coeliac disease. Trends Mol Med 16(11):537-50, 2010.

83. Plaza-Izurieta L, Castellanos-Rubio A, Irastorza I, Fernandez-Jimenez N, Gutierrez G, Bilbao JR. Revisiting genome wide association studies (GWAS) in coeliac disease: replication study in Spanish population and expression analysis of candidate genes. J Med Genet 48(7):493-6, 2011.

84. Cho JH. The genetics and immunopathogenesis of inflammatory bowel disease. Nat Rev Immunol 8(6):458-66, 2008.

85. Prasad S, Mingrino R, Kaukinen K, Hayes KL, Powell RM, MacDonald TT, et al. Inflammatory processes have differential effects on claudins 2, 3 and 4 in colonic epithelial cells. Lab Invest 85(9):1139-62, 2005.

86. Zeissig S, Burgel N, Gunzel D, Richter J, Mankertz J, Wahnschaffe U, et al. Changes in expression and distribution of claudin 2, 5 and 8 lead to discontinuous tight junctions and barrier dysfunction in active Crohn's disease. Gut 56(1):61-72, 2007.

87. Wehkamp J, Harder J, Weichenthal M, Schwab M, Schaffeler E, Schlee M, et al. NOD2 (CARD15) mutations in Crohn's disease are associated with diminished mucosal alpha-defensin expression. Gut 53(11):1658-64, 2004.

88. Wehkamp J, Salzman NH, Porter E, Nuding S, Weichenthal M, Petras RE, et al. Reduced Paneth cell alpha-defensins in ileal Crohn's disease. Proc Natl Acad Sci U S A 102(50):18129-34, 2005.

89. Wehkamp J, Schmid M, Stange EF. Defensins and other antimicrobial peptides in inflammatory bowel disease. Curr Opin Gastroenterol 23(4):370-8, 2007.

90. Ramasundara M, Leach ST, Lemberg DA, Day AS. Defensins and inflammation: the role of defensins in inflammatory bowel disease. J Gastroenterol Hepatol 24(2):202-8, 2009.

91. Elphick D, Liddell S, Mahida YR. Impaired luminal processing of human defensin-5 in Crohn's disease: persistence in a complex with chymotrypsinogen and trypsin. Am J Pathol 172(3):702-13, 2008.

92. Adolph TE, Tomczak MF, Niederreiter L, Ko HJ, Bock J, Martinez-Naves E, et al. Paneth cells as a site of origin for intestinal inflammation. Nature 503(7475):272-6, 2013.

93. Niess JH. Role of mucosal dendritic cells in inflammatory bowel disease. World J Gastroenterol 14(33):5138-48, 2008.

94. Bamias G, Sugawara K, Pagnini C, Cominelli F. The Th1 immune pathway as a therapeutic target in Crohn's disease. Curr Opin Investig Drugs 4(11):1279-86, 2003.

95. Targan SR, Karp LC. Defects in mucosal immunity leading to ulcerative colitis. Immunol Rev 206:296-305, 2005.

96. Ravi A, Garg P, Sitaraman SV. Matrix metalloproteinases in inflammatory bowel disease: boon or a bane? Inflamm Bowel Dis 13(1):97-107, 2007.

97. Fujino S, Andoh A, Bamba S, Ogawa A, Hata K, Araki Y, et al. Increased expression of interleukin 17 in inflammatory bowel disease. Gut 52(1):65-70, 2003.

98. Monteleone I, Sarra M, Pallone F, Monteleone G. Th17-related cytokines in inflammatory bowel diseases: friends or foes? Curr Mol Med 12(5):592-7, 2012.

99. Geremia A, Biancheri P, Allan P, Corazza GR, Di Sabatino A. Innate and adaptive immunity in inflammatory bowel disease. Autoimmun Rev 13(1):3-10, 2014.

100. Jabri B, Sollid LM. Tissue-mediated control of immunopathology in coeliac disease. Nat Rev Immunol 9(12):858-70, 2009.

101. Molberg O, McAdam SN, Korner R, Quarsten H, Kristiansen C, Madsen L, et al. Tissue transglutaminase selectively modifies gliadin peptides that are recognized by gut-derived T cells in celiac disease. Nat Med 4(6):713-7, 1998.

102. Spurkland A, Sollid LM, Polanco I, Vartdal F, Thorsby E. HLA-DR and -DQ genotypes of celiac disease patients serologically typed to be non-DR3 or non-DR5/7. Hum Immunol 35(3):188-92, 1992.

103. Sollid LM. Molecular basis of celiac disease. Annu Rev Immunol 18:53-81, 2000.

104. Briani C, Samaroo D, Alaedini A. Celiac disease: from gluten to autoimmunity. Autoimmun Rev 7(8):644-50, 2008.

105. Jabri B, Ebert E. Human CD8+ intraepithelial lymphocytes: a unique model to study the regulation of effector cytotoxic T lymphocytes in tissue. Immunol Rev 215:202-14, 2007.

106. Meresse B, Curran SA, Ciszewski C, Orbelyan G, Setty M, Bhagat G, et al. Reprogramming of CTLs into natural killer-like cells in celiac disease. J Exp Med 203(5):1343-55, 2006.

107. Meresse B, Chen Z, Ciszewski C, Tretiakova M, Bhagat G, Krausz TN, et al. Coordinated induction by IL15 of a TCR-independent NKG2D signaling pathway converts CTL into lymphokine-activated killer cells in celiac disease. Immunity 21(3):357-66, 2004.

108. Duerr RH, Taylor KD, Brant SR, Rioux JD, Silverberg MS, Daly MJ, et al. A genome-wide association study identifies IL23R as an inflammatory bowel disease gene. Science 314(5804):1461-3, 2006.

109. Barrett JC, Hansoul S, Nicolae DL, Cho JH, Duerr RH, Rioux JD, et al. Genome-wide association defines more than 30 distinct susceptibility loci for Crohn's disease. Nat Genet 40(8):955-62, 2008.

110. Olsen T, Rismo R, Cui G, Goll R, Christiansen I, Florholmen J. TH1 and TH17 interactions in untreated inflamed mucosa of inflammatory bowel disease, and their potential to mediate the inflammation. Cytokine 56(3):633-40, 2011.

111. Schmidt C, Giese T, Ludwig B, Mueller-Molaian I, Marth T, Zeuzem S, et al. Expression of interleukin-12-related cytokine transcripts in inflammatory bowel disease: elevated interleukin-23p19

and interleukin-27p28 in Crohn's disease but not in ulcerative colitis. Inflamm Bowel Dis 11(1):16-23, 2005.

112. Bogaert S, Laukens D, Peeters H, Melis L, Olievier K, Boon N, et al. Differential mucosal expression of Th17-related genes between the inflamed colon and ileum of patients with inflammatory bowel disease. BMC Immunol 11:61, 2010.

113. Fuss IJ, Becker C, Yang Z, Groden C, Hornung RL, Heller F, et al. Both IL-12p70 and IL-23 are synthesized during active Crohn's disease and are down-regulated by treatment with anti-IL-12 p40 monoclonal antibody. Inflamm Bowel Dis 12(1):9-15, 2006.

114. Castellanos-Rubio A, Santin I, Irastorza I, Castano L, Carlos Vitoria J, Ramon Bilbao J. TH17 (and TH1) signatures of intestinal biopsies of CD patients in response to gliadin. Autoimmunity 42(1):69-73, 2009.

115. Fernandez S, Molina IJ, Romero P, Gonzalez R, Pena J, Sanchez F, et al. Characterization of gliadin-specific Th17 cells from the mucosa of celiac disease patients. Am J Gastroenterol 106(3):528-38, 2011.

116. Monteleone I, Sarra M, Del Vecchio Blanco G, Paoluzi OA, Franze E, Fina D, et al. Characterization of IL-17A-producing cells in celiac disease mucosa. J Immunol 184(4):2211-8, 2010.

117. Bauquet AT, Jin H, Paterson AM, Mitsdoerffer M, Ho IC, Sharpe AH, et al. The costimulatory molecule ICOS regulates the expression of c-Maf and IL-21 in the development of follicular T helper cells and TH-17 cells. Nat Immunol 10(2):167-75, 2009.

118. Fina D, Sarra M, Fantini MC, Rizzo A, Caruso R, Caprioli F, et al. Regulation of gut inflammation and th17 cell response by interleukin-21. Gastroenterology 134(4):1038-48, 2008.

119. Yamanouchi J, Rainbow D, Serra P, Howlett S, Hunter K, Garner VE, et al. Interleukin-2 gene variation impairs regulatory T cell function and causes autoimmunity. Nat Genet 39(3):329-37, 2007.

120. Korn T, Oukka M, Kuchroo V, Bettelli E. Th17 cells: effector T cells with inflammatory properties. Semin Immunol 19(6):362-71, 2007.

121. Brand S. Crohn's disease: Th1, Th17 or both? The change of a paradigm: new immunological and genetic insights implicate Th17 cells in the pathogenesis of Crohn's disease. Gut 58(8):1152-67, 2009.

122. Benson JM, Peritt D, Scallon BJ, Heavner GA, Shealy DJ, Giles-Komar JM, et al. Discovery and mechanism of ustekinumab: a human monoclonal antibody targeting interleukin-12 and interleukin-23 for treatment of immune-mediated disorders. MAbs 3(6):535-45, 2011.

123. Sandborn WJ, Gasink C, Gao LL, Blank MA, Johanns J, Guzzo C, et al. Ustekinumab induction and maintenance therapy in refractory Crohn's disease. N Engl J Med 367(16):1519-28, 2012.

124. Toussirot E. The IL23/Th17 pathway as a therapeutic target in chronic inflammatory diseases. Inflamm Allergy Drug Targets 11(2):159-68, 2012.

125. Cosnes J, Cellier C, Viola S, Colombel JF, Michaud L, Sarles J, et al. Incidence of autoimmune diseases in celiac disease: protective effect of the gluten-free diet. Clin Gastroenterol Hepatol 6(7):753-8, 2008.

126. Lewis HM, Renaula TL, Garioch JJ, Leonard JN, Fry JS, Collin P, et al. Protective effect of gluten-free diet against development of lymphoma in dermatitis herpetiformis. Br J Dermatol 135(3):363-7, 1996.

127. Pricolo VE, Mangi AA, Aswad B, Bland KI. Gastrointestinal malignancies in patients with celiac sprue. Am J Surg 176(4):344-7, 1998.

128. Al-Toma A, Verbeek WH, Hadithi M, von Blomberg BM, Mulder CJ. Survival in refractory coeliac disease and enteropathy-associated T-cell lymphoma: retrospective evaluation of single-centre experience. Gut 56(10):1373-8, 2007.

129. Chaudhary R, Ghosh S. Prediction of response to infliximab in Crohn's disease. Dig Liver Dis 37(8):559-63, 2005.

130. Orlando A, Colombo E, Kohn A, Biancone L, Rizzello F, Viscido A, et al. Infliximab in the treatment of Crohn's disease: predictors of response in an Italian multicentric open study. Dig Liver Dis 37(8):577-83, 2005.

131. Parsi MA, Achkar JP, Richardson S, Katz J, Hammel JP, Lashner BA, et al. Predictors of response to infliximab in patients with Crohn's disease. Gastroenterology 123(3):707-13, 2002.

132. Arnott ID, McNeill G, Satsangi J. An analysis of factors influencing short-term and sustained response to infliximab treatment for Crohn's disease. Aliment Pharmacol Ther 17(12):1451-7, 2003.

133. Vermeire S, Louis E, Carbonez A, Van Assche G, Noman M, Belaiche J, et al. Demographic and clinical parameters influencing the short-term outcome of anti-tumor necrosis factor (infliximab) treatment in Crohn's disease. Serological markers for prediction of response to anti-tumor necrosis factor treatment in Crohn's disease. Am J Gastroenterol 97(9):2357-63, 2002.

134. Lionetti P, Bronzini F, Salvestrini C, Bascietto C, Canani RB, De Angelis GL, et al. Response to infliximab is related to disease duration in paediatric Crohn's disease. Aliment Pharmacol Ther 18(4):425-31, 2003.

135. Miheller P, Lakatos PL, Horvath G, Molnar T, Szamosi T, Czegledi Z, et al. Efficacy and safety of infliximab induction therapy in Crohn's Disease in Central Europe--a Hungarian nationwide observational study. BMC Gastroenterol 9:66, 2009.

136. Vermeire S, Louis E, Rutgeerts P, De Vos M, Van Gossum A, Belaiche J, et al. NOD2/CARD15 does not influence response to infliximab in Crohn's disease. Gastroenterology 123(1):106-11, 2002.

137. Mascheretti S, Hampe J, Croucher PJ, Nikolaus S, Andus T, Schubert S, et al. Response to infliximab treatment in Crohn's disease is

not associated with mutations in the CARD15 (NOD2) gene: an analysis in 534 patients from two multicenter, prospective GCP-level trials. Pharmacogenetics 12(7):509-15, 2002.

138. Mugnier B, Balandraud N, Darque A, Roudier C, Roudier J, Reviron D. Polymorphism at position -308 of the tumor necrosis factor alpha gene influences outcome of infliximab therapy in rheumatoid arthritis. Arthritis Rheum 48(7):1849-52, 2003.

139. Niewinski MM. Advances in celiac disease and gluten-free diet. J Am Diet Assoc 108(4):661-72, 2008.

140. Green PHR, Stavropoulos SN, Panagi SG, Goldstein SL, McMahon DJ, Absan H, et al. Characteristics of adult celiac disease in the USA: results of a national survey. Am J Gastroenterol 96(1):126-31, 2001.

141. D'Haens G, Baert F, van Assche G, Caenepeel P, Vergauwe P, Tuynman H, et al. Early combined immunosuppression or conventional management in patients with newly diagnosed Crohn's disease: an open randomised trial. Lancet 371(9613):660-7, 2008.

142. Abraham NS, Richardson P, Castillo D, Kane SV. Dual therapy with infliximab and immunomodulator reduces one-year rates of hospitalization and surgery among veterans with inflammatory bowel disease. Clin Gastroenterol Hepatol 11(10):1281-7, 2013.

143. Cleynen I, Gonzalez JR, Figueroa C, Franke A, McGovern D, Bortlik M, et al. Genetic factors conferring an increased susceptibility to develop Crohn's disease also influence disease phenotype: results from the IBDchip European Project. Gut, 62(11):1556-65, 2013.

144. Medrano LM, Dema B, Lopez-Larios A, Maluenda C, Bodas A, Lopez-Palacios N, et al. HLA and celiac disease susceptibility: new genetic factors bring open questions about the HLA influence and gene-dosage effects. PLoS One 7(10):e48403, 2012.

145. Bodd M, Kim CY, Lundin KE, Sollid LM. T-cell response to gluten in patients with HLA-DQ2.2 reveals requirement of peptide-MHC stability in celiac disease. Gastroenterology 142(3):552-61, 2012.

146. Kapoerchan VV, Wiesner M, Overhand M, van der Marel GA, Koning F, Overkleeft HS. Design of azidoproline containing gluten peptides to suppress CD4+ T-cell responses associated with celiac disease. Bioorg Med Chem 16(4):2053-62, 2008.

147. Brown G.J DJ, Marjason J.K, Ffrench R.A, Smith D, Sullivan M. A Phase I Study to Determine Safety, Tolerability and Bioactivity of Nexvax2 in HLA DQ2+ Volunteers With Celiac Disease Following a Long-Term, Strict Gluten-Free Diet. Gastroenterology 140(5):S-437-8, 2011.

148. Matoori S, Fuhrmann G, Leroux JC. Celiac disease: a challenging disease for pharmaceutical scientists. Pharm Res 30(3):619-26, 2013.

149. Medrano LMDB, Mendoza JL. Taxonera C, Díaz-Rubio M, G de la Concha E, Urcelay E, Núñez C. Interaction between TLR9 and IL23R polymorphisms influences Crohn´s disease susceptibility. Submitted, 2014.

150. Torok HP, Glas J, Endres I, Tonenchi L, Teshome MY, Wetzke M, et al. Epistasis between Toll-like receptor-9 polymorphisms and variants in NOD2 and IL23R modulates susceptibility to Crohn's disease. Am J Gastroenterol 104(7):1723-33, 2009.

151. Rumio C, Sommariva M, Sfondrini L, Palazzo M, Morelli D, Vigano L, et al. Induction of Paneth cell degranulation by orally administered Toll-like receptor ligands. J Cell Physiol 227(3):1107-13, 2012.

152. Hall JA, Bouladoux N, Sun CM, Wohlfert EA, Blank RB, Zhu Q, et al. Commensal DNA limits regulatory T cell conversion and is a natural adjuvant of intestinal immune responses. Immunity 29(4):637-49, 2008.

153. Mills KH. TLR9 turns the tide on Treg cells. Immunity 29(4):518-20, 2008.
154. Oukka M. Th17 cells in immunity and autoimmunity. Ann Rheum Dis 67 Suppl 3:iii26-9, 2008.
155. Troncone E, Marafini I, Pallone F, Monteleone G. Th17 Cytokines in Inflammatory Bowel Diseases: Discerning the Good from the Bad. Int Rev Immunol, 2013.
156. Hirota K, Yoshitomi H, Hashimoto M, Maeda S, Teradaira S, Sugimoto N, et al. Preferential recruitment of CCR6-expressing Th17 cells to inflamed joints via CCL20 in rheumatoid arthritis and its animal model. J Exp Med 204(12):2803-12, 2007.
157. Blauvelt A. T-helper 17 cells in psoriatic plaques and additional genetic links between IL-23 and psoriasis. J Invest Dermatol 128(5):1064-7, 2008.
158. Brennan FM, McInnes IB. Evidence that cytokines play a role in rheumatoid arthritis. J Clin Invest 118(11):3537-45, 2008.
159. Medrano LM, Garcia-Magarinos M, Dema B, Espino L, Maluenda C, Polanco I, et al. Th17-related genes and celiac disease susceptibility. PLoS One 7(2):e31244, 2012.
160. Cianci R, Cammarota G, Frisullo G, Pagliari D, Ianiro G, Martini M, et al. Tissue-infiltrating lymphocytes analysis reveals large modifications of the duodenal "immunological niche" in coeliac disease after gluten-free diet. Clin Transl Gastroenterol 3:e28, 2012.
161. Sjoberg V, Sandstrom O, Hedberg M, Hammarstrom S, Hernell O, Hammarstrom ML. Intestinal T-cell responses in celiac disease - impact of celiac disease associated bacteria. PLoS One 8(1):e53414, 2013.
162. Medrano L.M PV, Bodas A, Mendoza JL, Fernández-Arquero M, López-Palacios N, Núñez C. Th17 related gene expression in CeD and UC patients. Submitted, 2014.

163. Granlund A, Flatberg A, Ostvik AE, Drozdov I, Gustafsson BI, Kidd M, et al. Whole genome gene expression meta-analysis of inflammatory bowel disease colon mucosa demonstrates lack of major differences between Crohn's disease and ulcerative colitis. PLoS One 8(2):e56818, 2013.

164. Bi Y, Liu G, Yang R. Reciprocal modulation between TH17 and other helper T cell lineages. J Cell Physiol 226(1):8-13, 2011.

165. Zuniga LA, Jain R, Haines C, Cua DJ. Th17 cell development: from the cradle to the grave. Immunol Rev 252(1):78-88, 2013.

166. Matsukura H, Ikeda S, Yoshimura N, Takazoe M, Muramatsu M. Genetic polymorphisms of tumour necrosis factor receptor superfamily 1A and 1B affect responses to infliximab in Japanese patients with Crohn's disease. Aliment Pharmacol Ther 27(9):765-70, 2008.

167. Medrano LM, Taxonera C, Marquez A, Barreiro-de Acosta M, Gomez-Garcia M, Gonzalez-Artacho C, et al. Role of TNFRSF1B polymorphisms in the response of Crohn's disease patients to infliximab. Hum Immunol, 2013.

168. Arijs I, Quintens R, Van Lommel L, Van Steen K, De Hertogh G, Lemaire K, et al. Predictive value of epithelial gene expression profiles for response to infliximab in Crohn's disease. Inflamm Bowel Dis 16(12):2090-8, 2010.

169. Medrano LM, Taxonera C, Marquez A, Barreiro-de Acosta M, Gonzalez-Artacho C, Perez-Calle JL, et al. Validation of gene expression profile for response to infliximab in Crohn's disease. Submitted, 2014.

170. Craddock N, Hurles ME, Cardin N, Pearson RD, Plagnol V, Robson S, et al. Genome-wide association study of CNVs in 16,000 cases of eight common diseases and 3,000 shared controls. Nature 464(7289):713-20, 2010.

171. Pritchard JK. Are rare variants responsible for susceptibility to complex diseases? Am J Hum Genet 69(1):124-37, 2001.

172. Bamshad MJ, Ng SB, Bigham AW, Tabor HK, Emond MJ, Nickerson DA, et al. Exome sequencing as a tool for Mendelian disease gene discovery. Nat Rev Genet 12(11):745-55, 2011.

173. Rivas MA, Beaudoin M, Gardet A, Stevens C, Sharma Y, Zhang CK, et al. Deep resequencing of GWAS loci identifies independent rare variants associated with inflammatory bowel disease. Nat Genet 43(11):1066-73, 2011.

174. Abecasis GR, Altshuler D, Auton A, Brooks LD, Durbin RM, Gibbs RA, et al. A map of human genome variation from population-scale sequencing. Nature 467(7319):1061-73, 2010.

175. Huang J, Ellinghaus D, Franke A, Howie B, Li Y. 1000 Genomes-based imputation identifies novel and refined associations for the Wellcome Trust Case Control Consortium phase 1 Data. Eur J Hum Genet 20(7):801-5, 2012.

176. Ricano-Ponce I, Wijmenga C. Mapping of immune-mediated disease genes. Annu Rev Genomics Hum Genet 14:325-53, 2013.

177. Gamazon ER, Ziliak D, Im HK, LaCroix B, Park DS, Cox NJ, et al. Genetic architecture of microRNA expression: implications for the transcriptome and complex traits. Am J Hum Genet 90(6):1046-63, 2012.

178. Kumar V, Westra HJ, Karjalainen J, Zhernakova DV, Esko T, Hrdlickova B, et al. Human disease-associated genetic variation impacts large intergenic non-coding RNA expression. PLoS Genet 9(1):e1003201, 2013.

179. Festen EA, Goyette P, Green T, Boucher G, Beauchamp C, Trynka G, et al. A meta-analysis of genome-wide association scans identifies IL18RAP, PTPN2, TAGAP, and PUS10 as shared risk loci for Crohn's disease and celiac disease. PLoS Genet 7(1):e1001283, 2011.

180. Sawcer S, Hellenthal G, Pirinen M, Spencer CC, Patsopoulos NA, Moutsianas L, et al. Genetic risk and a primary role for cell-mediated immune mechanisms in multiple sclerosis. Nature 476(7359):214-9, 2011.

181. Li E, Hamm CM, Gulati AS, Sartor RB, Chen H, Wu X, et al. Inflammatory bowel diseases phenotype, C. difficile and NOD2 genotype are associated with shifts in human ileum associated microbial composition. PLoS One 7(6):e26284, 2012.

182. Wacklin P, Kaukinen K, Tuovinen E, Collin P, Lindfors K, Partanen J, et al. The duodenal microbiota composition of adult celiac disease patients is associated with the clinical manifestation of the disease. Inflamm Bowel Dis 19(5):934-41, 2013.

183. Nistal E, Caminero A, Herran AR, Arias L, Vivas S, de Morales JM, et al. Differences of small intestinal bacteria populations in adults and children with/without celiac disease: effect of age, gluten diet, and disease. Inflamm Bowel Dis 18(4):649-56, 2012.

184. Sanz Y, De Pama G, Laparra M. Unraveling the ties between celiac disease and intestinal microbiota. Int Rev Immunol 30(4):207-18, 2011.

## *Anexo I: Artículos originales*

OPEN ACCESS Freely available online

PLOS ONE

# HLA and Celiac Disease Susceptibility: New Genetic Factors Bring Open Questions about the HLA Influence and Gene-Dosage Effects

**Luz María Medrano[1], Bárbara Dema[1], Arturo López-Larios[1], Carlos Maluenda[2], Andrés Bodas[2], Natalia López-Palacios[3], M. Ángeles Figueredo[1], Miguel Fernández-Arquero[1], Concepción Núñez[1]***

**1** UGC de Inmunología, Hospital Clínico San Carlos, Instituto de Investigación Sanitaria del Hospital Clínico San Carlos (IdISSC), Madrid, Spain, **2** Servicio de Pediatría, Hospital Clínico San Carlos, Instituto de Investigación Sanitaria del Hospital Clínico San Carlos (IdISSC), Madrid, Spain, **3** Servicio de Aparato Digestivo, Hospital Clínico San Carlos, Instituto de Investigación Sanitaria del Hospital Clínico San Carlos (IdISSC), Madrid, Spain

## Abstract

Celiac disease (CD) is a chronic inflammatory disorder triggered after gluten ingestion in genetically susceptible individuals. The major genetic determinants are *HLA-DQA1*05* and *HLA-DQB1*02*, which encode the DQ2 heterodimer. These alleles are commonly inherited in cis with *DRB1*03:01*, which is associated with numerous immune-related disorders, in some cases contributing with a different amount of risk depending on the haplotype context. We aimed at investigating those possible differences involving *DRB1*03:01*-carrying haplotypes in CD susceptibility. A family (274 trios) and a case-control sample (369 CD cases/461 controls) were analyzed. *DRB1*03:01*-carrying individuals were classified according to the haplotype present (ancestral haplotype (AH) 8.1, AH 18.2 or non-conserved haplotype) after genotyping of *HLA-DRB1*, *-DQA1*, *-DQB1*, *-B8*, *TNF* -308, *TNF* -376 and the *TNFa* and *TNFb* microsatellites. We observe that the AH 8.1 confers higher risk than the remaining *DRB1*03:01*-carrying haplotypes, and this effect only involves individuals possessing a single copy of *DQB1*02*. CD risk for these individuals is similar to the one conferred by inherit *DQA1*05* and *DQB1*02* in trans. It seems that an additional CD susceptibility factor is present in the AH 8.1 but not in other *DRB1*03:01*-carrying haplotypes. This factor could be shared with individuals possessing DQ2.5 trans, according to the similar risk observed in those two groups of individuals.

**Citation:** Medrano LM, Dema B, López-Larios A, Maluenda C, Bodas A, et al. (2012) HLA and Celiac Disease Susceptibility: New Genetic Factors Bring Open Questions about the HLA Influence and Gene-Dosage Effects. PLoS ONE 7(10): e48403. doi:10.1371/journal.pone.0048403

**Editor:** Massimo Pietropaolo, University of Michigan Medical School, United States of America

**Received** July 30, 2012; **Accepted** October 1, 2012; **Published** October 31, 2012



**Funding:** This work was supported by grants from Fondo de Investigaciones Sanitarias (PI11/00614 and PI10/02876). The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

**Competing Interests:** The authors have declared that no competing interests exist.

* E-mail: conchita.npardo@gmail.com

## Introduction

Human leukocyte antigen (HLA) is a master piece in the pathogenesis of celiac disease (CD), as first evidenced by the strong genetic association existent between CD susceptibility and certain HLA alleles. This region, located on 6p21, contains hundreds of genes with immunological function and it is characterized by a high gene density and variability and an extensive linkage disequilibrium, which make difficult to pinpoint the causal variant/s. Despite this, CD can be considered a particular disease since the specific HLA alleles involved and their functional implication are well-established. The presence of the *DQA1*05* and *DQB1*02* susceptibility alleles implies the formation of the $\alpha$ and $\beta$ chains of the HLA-DQ2 heterodimer, present in around 90–95% of CD individuals. This molecule shows high affinity for peptides resultant from incomplete gluten digestion, which bind and present to antigen specific T cells, triggering the intestinal inflammation prototypical of CD.

In most cases, *DQA1*05* and *DQB1*02* are encoded in the same chromosome (DQ2.5 cis) and appear in very strong linkage disequilibrium with *DRB1*03:01*. In fact, this allele was first associated with CD risk [1]. The *DQA1*05*, *DQB1*02* and *DRB1*03:01* alleles can be present in two different haplo-specific contexts and constitute the so-called ancestral haplotypes (AH) 8.1 and AH 18.2; they can also be found within non-specific allelic combinations and constitute other less frequent haplotypes (hereafter called non-conserved haplotypes). The *DRB1*03:01* allele, and consequently, *DRB1*03:01* haplotypes, have been associated to numerous immune-mediated disorders, as type 1 diabetes, multiple sclerosis or selective IgA deficiency, among many others; in some cases with the different *DRB1*03:01* haplotypes showing a differential contribution to disease risk [2,3]. A differential behaviour between AH 18.2 and AH 8.1 has also been described in CD [4,5], although no relevance has been given to this observation. The *DQA1*05* and *DQB1*02* alleles can also be inherited in trans, each encoded in one chromosome from each parent (DQ2.5 trans).

HLA influence on CD susceptibility shows a dose effect. Individuals can be classified in high or intermediate CD risk according to the number of *DQA1*05*- and *DQB1*02*-carrying alleles. Homozigosity for DQ2.5 cis and heterozigosity for DQ2.5 cis with a chromosome possessing a second *DQB1*02* allele (DQ2.2) confer the highest risk to develop CD. Heterozigosity for DQ2.5 cis in individuals with a single copy of *DQB1*02* (non-DQ2.2) or presence of DQ2.5 trans confer intermediate risk.

Additionally to the molecule DQ2.5, the influence of HLA-DQ8 (genetically *DQA1*03*, *DQB1*03:02*) on the disease is already known. This molecule is present in almost all the CD patients without DQ2.5. However, the genetic influence of the HLA region in CD is not limited to the factors coding DQ2 or DQ8, and several works have attempted to discover new susceptibility factors without much success (see [6] for review). Some variants in the *TNF* gene have been suggested as DQ2 independent factors for CD susceptibility, even as the responsible factors for the additional risk present on the AH 8.1. [7]. Last years have witnessed a spectacular increase in the knowledge of the genetic basis of CD, favoured by development of genome wide association studies (GWAS), but these works have not added new information about the HLA contribution because they have been mainly focused on the influence of genes outside this region.

We aimed at investigating the additional genetic contribution to CD susceptibility lying on the HLA region, by focusing in the possible differential contribution of the different *DRB1*03:01*-carrying haplotypes.

## Materials and Methods

### Ethics Statement

This study was approved by the ethical committee (CEIC) of Hospital Clínico San Carlos. Samples were obtained after obtaining written informed consent. For children, the informed consent was signed by their parents or legal guardian.

### Subjects

A total of 274 trios composed for both parents and the affected child and a case-control series consisting of 369 independent CD patients and 461 ethnically matched healthy controls were studied. CD patients were diagnosed following the European Society for Pediatric Gastroenterology, Hepatology, and Nutrition (ESPGHAN) [8], 97% are positive for HLA-DQ2 and/or HLA-DQ8. Controls correspond mainly to blood donors and laboratory staff. CD samples were consecutively collected in two centres of the same region (Hospital La Paz and Hospital Clínico San Carlos, Madrid, Spain) and controls were collected at the Hospital Clínico San Carlos. All samples correspond to unrelated Spanish white individuals.

### Genotyping

DNA was extracted from fresh peripheral blood leukocytes by a "salting out" procedure. All samples were genotyped for *HLA-DRB1*, *-DQA1* and *-DQB1* by PCR-SSOP (Polymerase Chain Reaction-Sequence Specific Oligonucleotide Probe). The different *DRB1*03:01* haplotypes were assessed by additional genotyping of the *TNF* single nucleotide polymorphisms (SNPs) -308 (rs1800629) and -376 (rs1800750) and the microsatellites *TNFa* and *TNFb*; those polymorphisms were typed as previously described [9,10]. The presence of the *HLA-B8* allele was tested by TaqMan technology using the tag SNPs rs6457374 and rs2844535 (Applied Biosystems Inc., Foster City, CA, USA).

We defined the AH 8.1 according to the simultaneously presence of *DRB1*03:01*, *DQA1*05:01*, *DQB1*02:01*, *TNF -308A*, *TNFa2b3* and *B*8*. AH 18.2 was defined according to carriage of *DRB1*03:01*, *DQA1*05:01*, *DQB1*02:01*, *TNF -376A* and *TNFa1b5*. Haplotypes with all the remaining allelic combinations in those loci or markers were designed as non-conserved haplotypes. The *TNF* markers studied were selected because, in *DRB1*03:01* subjects, they are haplo-specific for AH 18.2 or AH 8.1.

### Statistical Analysis

HLA haplotypes were deduced directly from the pedigree for patients used in the family study. In cases and controls, the EM (Expectation-Maximization) algorithm implemented in the Arlequin software was used to estimate haplotype frequencies.

The transmission disequilibrium test (TDT) was used to analyse the preferential transmission of one haplotype over the others when analysing family data. This test uses only information provided by heterozygous parents.

Comparisons between groups were performed with the chi-square test using the statistical package EpiInfo v5.00 (CDC, Atlanta, USA). Heterogeneity between haplotype groups was evaluated with Review Manager (RevMan) 5.0 software (Copenhagen: The Nordic Cochrane Centre, The Cochrane Collaboration, 2008).

## Results

We studied 274 trios to investigate the possibility of a differential transmission of the different *DRB1*03:01* haplotypes to the affected child (Table 1). *DRB1*03:01* is always preferentially transmitted, independently of its haplotype context (see TDT results, in Table 1). However, the distortion in the transmission of this allele is significantly higher when it is present in the AH 8.1, compared with its presence in the remaining *DRB1*03:01*-containing haplotypes ($p = 8.7*10^{-4}$ vs. AH 18.2; $p = 2.4*10^{-4}$ vs. non-conserved haplotypes). AH 18.2 and non-conserved haplotypes show a similar preferential transmission to offspring ($p = 0.99$). These differences are also observed when considering the haplotype transmission from the *DRB1*03:01* homozygous parents (composed by different *DRB1*03:01* haplotypes) included in the 274 families (Table 2).

We wanted to validate this observation in a case-control sample (Table 3). Since no differences were observed between AH 18.2 and non-conserved haplotypes in the family data, they were combined in subsequent analysis (and called non-AH 8.1). As already known, *DRB1*03:01* overall appears at significantly higher frequency in CD patients than in controls: 45% vs. 14%, respectively (OR = 4.97 95% CI 3.90–6.34, $p<10^{-7}$); but we additionally show that this case-control difference is higher when considering only the AH 8.1 (OR = 6.53 95% CI 4.47–9.56, $p<10^{-7}$).

**Table 1.** Transmission data of the diverse *DRB1*03:01*-containing haplotypes in the 274 families studied and TDT results.

| DRB1*03:01 HAPLOTYPES | T | U | TDT |
|---|---|---|---|
| 1. All | 268 | 88 | $p<10^{-10}$ |
| 2. AH 8.1 | 118 | 18 | $p<10^{-10}$ |
| 3. AH 18.2 | 58 | 27 | $p = 2.2*10^{-4}$ |
| 4. Non-conserved | 92 | 41 | $p = 4.9*10^{-6}$ |
| 5. Non-AH 8.1 | 150 | 68 | $p = 5.3*10^{-9}$ |

T = transmitted; U = untransmitted; TDT = transmission disequilibrium test.
Non-AH 8.1 includes AH18.2 and non-conserved haplotypes.
TDT results are calculated after excluding homozygous parents (5 for AH 8.1, 5 for AH 18.2 and 7 for *DRB1*03:01* non-conserved haplotypes).
Haplotype comparisons: 2 vs. 3: p = 0.00087 OR = 3.05 95% CI 1.48–6.33; 2 vs. 4: p = 0.00024 OR = 3.06 95% CI 1.59–5.94; 3 vs. 4: p = 0.99 OR = 1.00 95% CI 0.54–1.88; 2 vs. 5: $p = 7.8*10^{-5}$ OR = 3.06 95% CI 1.67–5.65.
doi:10.1371/journal.pone.0048403.t001

**Table 2.** Transmission data of the diverse *DRB1*03:01*-containing haplotypes from *DRB1*03:01* homozygous parents (N = 27).

| HAPLOTYPE COMPOSITION | N | T | TDT |
|---|---|---|---|
| AH 8.1 - AH 18.2 | 6 | 6 AH 8.1, 0 AH 18.2 | p = 0.016 |
| AH 8.1 - Non-conserved | 10 | 7 AH 8.1, 3 Non-conserved | p = 0.17 |
| AH 18.2 - Non-conserved | 11 | 4 AH 18.2, 7 Non-conserved | p = 0.27 |
| AH 8.1– Non-AH 8.1 | 16 | 13 AH 8.1, 3 Non- AH 8.1 | p = 0.011 |

T = transmitted; TDT = transmission disequilibrium test.
doi:10.1371/journal.pone.0048403.t002

In CD, it is well established the existence of a dose effect, what implies the existence of different CD risk categories attending to their HLA constitution. We investigated this differential risk contribution of *DRB1*03:01*-containing haplotypes in those categories and found that only in those individuals carrying a single copy of the *DQB1*02* allele (individuals DQ2.5 cis + non-DQ2.2 in Table 3), the presence of AH 8.1 confers additional risk.

When considering the HLA risk categories according to gene dosage effects, carriage of the DQ2 molecule in individuals with a single copy of the *DQB1*02* allele is considered as conferring intermediate CD risk. A similar risk is conferred by the presence of DQ2.5 trans, although some groups reported this to be an intermediate higher risk group [11]. We compared CD risk in carriers of DQ2.5 trans (48 individuals out of 369 patients and 21 out of 461 controls) to CD risk in carriers of DQ2.5 cis with and without the AH 8.1 and we observed that DQ2.5 trans confers similar risk than DQ2.5 cis with AH 8.1 (heterogeneity: p = 0.91, $I^2 = 0\%$) and significantly higher risk than DQ2.5 cis with non-AH 8.1 haplotypes (heterogeneity: p = 0.09, $I^2 = 64\%$).

Finally, we investigated the possibility that the similar risk conferred by DQ2.5 trans and DQ2.5 cis with AH 8.1 was due to the presence of a common susceptibility factor. In most cases, carriers of the molecule DQ2.5 trans are genetically characterized by being heterozygous *DQB1*03:01-DQA1*05:05/DQB1*02:02-DQA1*02:01* (serologically DR5/DR7, terms also used hereafter for simplification purposes). We used genotype data corresponding to 6,769 SNPs located in the HLA (29.96–33.19 Mb interval), which were previously obtained in a subset of our Spanish samples (more than 500 CD patients and 300 controls) in the context of the Immunochip Project (http://www.immunobase.org). We selected all the homozygous individuals for AH 8.1, AH 18.2, DR5 (*DQB1*03:01-DQA1*05:05*) and DR7 (*DQB1*02:02-DQA1*02:01*) and looked for a chromosomal region shared between AH 8.1 and DR5 (*DQB1*03:01-DQA1*05:05*) or AH 8.1 and DR7 (*DQB1*02:02-DQA1*02:01*) but not common to AH 18.2. However, no regions emerged after that search. Note that this result could be affected by the low number of available homozygous individuals, mainly DR5 (*DQB1*03:01-DQA1*05:05*) and DR7 (*DQB1*02:02-DQA1*02:01*) (3 and 4, respectively).

## Discussion

Our results evidence the presence of an additional susceptibility factor to CD in the HLA region, which is linked to AH 8.1. This factor only increases susceptibility when appearing in individuals carrying a single copy of the *DQB1*02* allele (DQ2.5 cis non-DQ2.2) and split up the HLA intermediate risk group into two groups: one with higher intermediate risk, which is composed of

**Table 3.** Frequency and comparison of the diverse *DRB1*03:01* haplotypes in celiac disease (CD) patients (2N = 738) and controls (2N = 922), overall and classified according to the different HLA CD risk categories.

| DRB1*03:01 HAPLOTYPES | CD | | CONTROLS | | CD vs. CONTROLS | | HETEROGENEITY AH 8.1 VS. NON-AH 8.1 | |
|---|---|---|---|---|---|---|---|---|
| | 2N | % | 2N | % | p | OR (95% CI) | p | $I^2$ |
| **ALL DRB1*03:01** | | | | | | | | |
| All | 330 | 44.7 | 129 | 14.0 | $<10^{-7}$ | 4.97 (3.90–6.34) | | |
| AH 8.1 | 141 | 19.1 | 42 | 4.6 | $<10^{-7}$ | 6.53 (4.47–9.56) | 0.06 | 71% |
| Non-AH 8.1 | 189 | 25.6 | 87 | 9.4 | $<10^{-7}$ | 4.22 (3.16–5.65) | | |
| **DQ2.5 cis+DQ2.5 cis** | | | | | | | | |
| All | 74 | 10.0 | 16 | 1.7 | $<10^{-7}$ | 8.99 (5.03–16.28) | | |
| AH 8.1 | 30 | 4.1 | 6 | 0.65 | $<10^{-7}$ | 9.72 (3.93–28.74) | 0.82 | 0% |
| Non- AH 8.1 | 44 | 6.0 | 10 | 1.1 | $<10^{-7}$ | 8.55 (4.10–18.31) | | |
| **DQ2.5 cis+DQ2.2** | | | | | | | | |
| All | 104 | 14.1 | 13 | 1.4 | $<10^{-7}$ | 15.55 (8.39–29.37) | | |
| AH 8.1 | 40 | 5.4 | 6 | 0.7 | $<10^{-7}$ | 12.96 (5.39–37.64) | 0.60 | 0% |
| Non- AH 8.1 | 64 | 8.7 | 7 | 0.8 | $<10^{-7}$ | 17.77 (8.03–46.30) | | |
| **DQ2.5 cis+non-DQ2.2** | | | | | | | | |
| All | 152 | 20.6 | 100 | 10.8 | $<10^{-7}$ | 2.95 (2.21–3.94) | | |
| AH 8.1 | 72 | 9.8 | 30 | 3.3 | $<10^{-7}$ | 4.66 (2.94–7.44) | 0.009 | 85% |
| Non-AH 8.1 | 80 | 10.8 | 70 | 7.6 | $3.2*10^{-6}$ | 2.22 (1.56–3.17) | | |

% are referred to the total number of haplotypes.
Non-AH 8.1 includes AH 18.2 and non-conserved haplotypes.
DQ2.5 cis = *DQB1*02:01-DQA1*05:01*; DQ2.2 = *DQB1*02:02-DQA1*02:01*.
doi:10.1371/journal.pone.0048403.t003

carriers of the molecule DQ2.5 cis encoded by AH 8.1 and carriers of the molecule DQ2.5 trans; and one with lower intermediate risk, which is composed of carriers of the molecule DQ2.5 cis encoded by either AH 18.2 or non-conserved *DRB1*03:01* haplotypes.

The well-known contribution of the DQ2 molecule to CD pathogenesis is genetically based on carriage of *DQA1*05* and *DQB1*02*. These two alleles are commonly present in the small segment identical by descent among *DRB1*03:01*-containing haplotypes. However, outside that shared segment, the divergence between *DRB1*03:01* haplotypes do not differ from that found between disparate haplotypes [12]; therefore, a susceptibility factor located there is not expected to be present in all *DRB1*03:01* haplotypes. On the other hand, previous studies suggested a close evolutionary relationship among *DRB1*03:01*-containing haplotypes, DR5 and DR7 (*DQB1*03:01-DQA1*05:05* and *DQB1*02:02-DQA1*02:01*) [13], which could explain the existence of a common susceptibility factor between AH 8.1 and one of those haplotypes. No definitive conclusion can de drawn from our data due to the low number of chromosomes compared. Further analysis including higher sample size is mandatory.

The analysis of the HLA region in GWAS showed, besides the expected peak corresponding to DQ2.5 cis, two SNPs associated to CD risk, both located within or adjacent to *HLA-DQA1* and *HLA-DQB1* [14]. This is in accordance with our results, which suggest a risk factor common to *DRB1*03:01*-containing haplotypes and DR5 or DR7 (*DQB1*03:01-DQA1*05:05* or *DQB1*02:02-DQA1*02:01*), because those are the regions that they shared.

One intriguing issue emerging from our study is why the additional risk factor present in the AH 8.1 does not seem to influence on CD susceptibility when it appears in individuals carrying a second copy of the *DQB1*02* allele. In CD, T cell stimulation due to gluten-derived peptides depends on the number and type of HLA-DQ2 molecules expressed. DQ2.5 molecules can bind a high repertoire of gluten peptides, but only a restricted subset is bound to DQ2.2 molecules, which reduce the immunogenicity of DQ2.2. Additionally, the number of these DQ molecules is also a relevant factor in T cell stimulation and this depends on the number of specific alleles in *DQA1* and *DQB1* loci, which determines the possible αβ-chain combinations constituting the DQ heterodimers [15]. As a matter of fact, all HLA-DQ molecules are identical in HLA-DQ2.5 homozygous individuals, which can bind a very high repertoire of gluten peptides and confer the highest CD risk. It could be speculated that in such scenario an additional susceptibility factor has null or limited possibility to increase risk. By the contrary, in HLA-DQ2.5 cis individuals (without a second copy of *DQB1*02, i.e.*, non-DQ2.2), only one of the four possible αβ combinations constitutes an HLA-DQ2.5 molecule and therefore the presence of a genetic factor which increases immunogenicity against gluten derived peptides could have a relevant impact in increasing CD risk.

The HLA dose effect is also influenced by differences in the kinetic stability of the interaction between HLA molecules and gluten derived peptides, key factor for development of T cell responses against gluten [16,17]. For most peptide ligands, DQ2.5 shows higher binding stability than DQ2.2. The risk variant present in the AH 8.1, and presumably in individuals possessing DQ2.5 trans molecules, could affect the kinetic stability of the interaction HLA-gluten either increasing the number of gluten derived peptides which can bind with high affinity or increasing the binding stability of peptides previously recognised. Bodd *et al* [17] claimed that T-cell epitopes must be assessed and characterized in the context of the HLA molecules expressed by the T-cell donor and underlined the relevance that this could have for future peptide-based vaccines. According to that, it would be interesting to establish a comparison of the HLA-DQ molecules present in individuals DQ2.5 cis with AH 8.1, DQ2.5 cis with non-AH 8.1 haplotypes and DQ2.5 trans. Differences in their binding gluten peptides would imply that peptide-based vaccines should look at those individuals differentially.

Much more work deserves this field, with several open questions as which is the causal variant lying on the AH 8.1 and which are their specific functional implications.

## Acknowledgments

We are most grateful to Carmen Martínez Cuervo and M. Ángel García Martínez for their expert technical assistance.

## Author Contributions

Conceived and designed the experiments: CN. Performed the experiments: LMM BD ALL. Analyzed the data: LMM BD ALL MFA CN. Contributed reagents/materials/analysis tools: CM AD AB NLP MAF. Wrote the paper: LMM CN. All authors critically revised the manuscript and approved the final version of the manuscript.

## References

1. Keuning JJ, Pena AS, van Leeuwen A, van Hooff JP, va Rood JJ (1976) HLA-DW3 associated with coeliac disease. Lancet 1: 506–508.
2. Baschal EE, Aly TA, Jasinski JM, Steck AK, Noble JA, et al. (2009) Defining multiple common "completely" conserved major histocompatibility complex SNP haplotypes. Clin Immunol 132: 203–214.
3. de la Concha EG, Cavanillas ML, Cenit MC, Urcelay E, Arroyo R, et al. (2012) DRB1*03:01 Haplotypes: Differential Contribution to Multiple Sclerosis Risk and Specific Association with the Presence of Intrathecal IgM Bands. PLoS One 7: e31018.
4. Bolognesi E, Karell K, Percopo S, Coto I, Greco L, et al. (2003) Additional factor in some HLA DR3/DQ2 haplotypes confers a fourfold increased genetic risk of celiac disease. Tissue Antigens 61: 308–316.
5. Bilbao JR, Calvo B, Aransay AM, Martin-Pagola A, Perez de Nanclares G, et al. (2006) Conserved extended haplotypes discriminate HLA-DR3-homozygous Basque patients with type 1 diabetes mellitus and celiac disease. Genes Immun 7: 550–554.
6. Louka AS, Lie BA, Talseth B, Ascher H, Ek J, et al. (2003) Coeliac disease patients carry conserved HLA-DR3-DQ2 haplotypes revealed by association of TNF alleles. Immunogenetics 55: 339–343.
7. de la Concha EG, Fernandez-Arquero M, Vigil P, Rubio A, Maluenda C, et al. (2000) Celiac disease and TNF promoter polymorphisms. Hum Immunol 61: 513–517.
8. Husby S, Koletzko S, Korponay-Szabo IR, Mearin ML, Phillips A, et al. (2012) European Society for Pediatric Gastroenterology, Hepatology, and Nutrition guidelines for the diagnosis of coeliac disease. J Pediatr Gastroenterol Nutr 54: 136–160.
9. Nedospasov SA, Udalova IA, Kuprash DV, Turetskaya RL (1991) DNA sequence polymorphism at the human tumor necrosis factor (TNF) locus. Numerous TNF/lymphotoxin alleles tagged by two closely linked microsatellites in the upstream region of the lymphotoxin (TNF-beta) gene. J Immunol 147: 1053–1059.
10. Fernandez-Arquero M, Arroyo R, Rubio A, Martin C, Vigil P, et al. (1999) Primary association of a TNF gene polymorphism with susceptibility to multiple sclerosis. Neurology 53: 1361–1363.
11. van Heel DA, Hunt K, Greco L, Wijmenga C (2005) Genetics in coeliac disease. Best Pract Res Clin Gastroenterol 19: 323–339.
12. Traherne JA, Horton R, Roberts AN, Miretti MM, Hurles ME, et al. (2006) Genetic analysis of completely sequenced disease-associated MHC haplotypes identifies shuffling of segments in recent human history. PLoS Genet 2: e9.
13. Sollid LM (2000) Molecular basis of celiac disease. Annu Rev Immunol 18: 53–81.
14. van Heel DA, Franke L, Hunt KA, Gwilliam R, Zhernakova A, et al. (2007) A genome-wide association study for celiac disease identifies risk variants in the region harboring IL2 and IL21. Nat Genet 39: 827–829.
15. Vader W, Stepniak D, Kooy Y, Mearin L, Thompson A, et al. (2003) The HLA-DQ2 gene dose effect in celiac disease is directly related to the magnitude and breadth of gluten-specific T cell responses. Proc Natl Acad Sci U S A 100: 12390–12395.

16. Fallang LE, Bergseng E, Hotta K, Berg-Larsen A, Kim CY, et al. (2009) Differences in the risk of celiac disease associated with HLA-DQ2.5 or HLA-DQ2.2 are related to sustained gluten antigen presentation. Nat Immunol 10: 1096–1101.

17. Bodd M, Kim CY, Lundin KE, Sollid LM (2012) T-Cell Response to Gluten in Patients With HLA-DQ2.2 Reveals Requirement of Peptide-MHC Stability in Celiac Disease. Gastroenterology 142: 552–561.

OPEN ACCESS Freely available online

PLoS one

# Th17-Related Genes and Celiac Disease Susceptibility

**Luz María Medrano[1], Manuel García-Magariños[2,3], Bárbara Dema[1], Laura Espino[1], Carlos Maluenda[4], Isabel Polanco[5], M. Ángeles Figueredo[1], Miguel Fernández-Arquero[1], Concepción Núñez[1]***

**1** UGC de Inmunología, Hospital Clínico San Carlos, Instituto de Investigación Sanitaria del Hospital Clínico San Carlos (IdISSC), Madrid, Spain, **2** Unidade de Xenética, Instituto de Medicina Legal and Departamento de Anatomía Patológica y Ciencias Forenses, Facultade de Medicina, Universidade de Santiago de Compostela, Santiago de Compostela, Spain, **3** Departamento de Estadística e IO, Universidad Pública de Navarra, Pamplona, Spain, **4** Servicio de Pediatría, Hospital Clínico San Carlos, Madrid, Spain, **5** Servicio de Gastroenterología Pediátrica, Hospital La Paz, Madrid, Spain

## Abstract

Th17 cells are known to be involved in several autoimmune or inflammatory diseases. In celiac disease (CD), recent studies suggest an implication of those cells in disease pathogenesis. We aimed at studying the role of genes relevant for the Th17 immune response in CD susceptibility. A total of 101 single nucleotide polymorphisms (SNPs), mainly selected to cover most of the variability present in 16 Th17-related genes (*IL23R, RORC*, *IL6R*, *IL17A*, *IL17F*, *CCR6*, *IL6*, *JAK2*, *TNFSF15*, *IL23A*, *IL22*, *STAT3*, *TBX21*, *SOCS3*, *IL12RB1* and *IL17RA*), were genotyped in 735 CD patients and 549 ethnically matched healthy controls. Case-control comparisons for each SNP and for the haplotypes resulting from the SNPs studied in each gene were performed using chi-square tests. Gene-gene interactions were also evaluated following different methodological approaches. No significant results emerged after performing the appropriate statistical corrections. Our results seem to discard a relevant role of Th17 cells on CD risk.

**Citation:** Medrano LM, García-Magariños M, Dema B, Espino L, Maluenda C, et al. (2012) Th17-Related Genes and Celiac Disease Susceptibility. PLoS ONE 7(2): e31244. doi:10.1371/journal.pone.0031244

**Editor:** Aftab A. Ansari, Emory University School of Medicine, United States of America

**Received** October 28, 2011; **Accepted** January 5, 2012; **Published** February 16, 2012



**Funding:** This work was supported by a grant from "Fondo de Investigaciones Sanitarias" (CP08/00213). The funder had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

**Competing Interests:** The authors have declared that no competing interests exist.

* E-mail: conchita.npardo@gmail.com

## Introduction

Celiac disease (CD) is an immune related disease mainly characterized by intestinal inflammation after gluten ingestion in genetically susceptible individuals. CD has been traditionally considered a Th1-mediated disease. However, accumulating evidence about the relevant role of the novel Th17 immune response in several autoimmune diseases [1] opened the possibility towards an involvement of this immunological pathway in CD pathogenesis. These cells seem to be involved in protective responses against extracellular pathogens but they can contribute to chronic inflammation and autoimmunity when dysregulated.

Th17 cells develop from naïve CD161+ CD4+ T cells upon stimulation with particular immunological stimulus, especifically, transforming growth factor beta (TGF-β), interleukin (IL)-23, IL-1β or IL-6 [2]. This induces several transcription factors, mainly the RAR-related orphan receptor C (RORC), which in turn activates IL-17A and IL-17F transcription, the distinctive effector cytokines of this subset of T cells. Production of IL-21, IL-22 and IL-26 also characterizes this specific response, besides the surface markers C-C chemokine receptor type 6 (CCR6) and IL-23 receptor (IL-23R).

Studies based on murine models of several autoimmune diseases, as multiple sclerosis (autoimmune encephalomyelitis, EAE), rheumatoid arthritis (collagen-induced arthritis, CIA) and inflammatory bowel disease (experimental colitis), provided the first evidence about a role of Th17 cells in those conditions [3,4]. This idea was later supported by case-control studies, which associated genetic variants in *IL23R* with susceptibility to Crohn's disease, psoriasis and ankylosing spondylitis [5,6,7]. Nowadays, the Th17 immune response is considered as a relevant player in several autoimmune or inflammatory diseases. IL-17 mRNA or protein have been detected in biological fluids or the specific affected tissue in several autoimmune disorders [8] and genetic studies associated genes coding important Th17 related products with several diseases [9]. In addition, epistasis between *IL23R* and other Th17 related genes has been reported: with *IL2/IL21* in UC [10] and with *IL17A* and *IL17RA* in Crohn's disease [11].

In 2008, a putative implication of the Th17 immune response in CD pathogenesis was suggested from two studies following different approaches. Our research group detected a significant association between a genetic polymorphism in the *IL23R* gene and CD [12] and Harris et al. found higher production of IL-23 after stimulation of human monocytes derived from CD patients with peptic fragments of wheat gliadin [13]. Subsequently, genetic linkage with the *IL23R* region was observed in Finnish families, although this result was not replicated in Hungarian pedigrees and no association with *IL23R* polymorphisms was observed in Finnish, Hungarian or Italian CD samples [14]. In addition, increased expression of several Th17-related cytokines or products was detected in patients with active CD [15,16] and very recently, gluten-specific IL-17A-producing cells have been found in the duodenum of CD patients [17], which supports a role of Th17 cells in CD pathogenesis.

Despite these results observed in CD, the role of the Th17 cells on this disease is still not well defined. We aimed at shedding more light upon this issue by performing an extensive genetic study including many genes coding distinctive cytokines, markers or

transcription factors involved in the Th17 response. We will evaluate the individual influence of those genes on CD susceptibility and also the possible contribution of gene-gene interactions. Previous genome wide association studies (GWAS) did not find association with CD susceptibility of any Th17-related gene [18,19,20] (with exception of the *IL2/IL21* locus, also involved in other processes), but we consider that a different scenario could emerge with this study: we cover most of the variability present in the studied genetic regions and we will evaluate the genetic interactions between the included polymorphisms, which has been proved as a valid approach to detect new susceptibility variants [21,22].

## Materials and Methods

### Ethics Statement

This study was approved by the ethical committee (CEIC) of the Hospital Clínico San Carlos. Samples were obtained after obtaining written informed consent.

### Subjects

A total of 735 CD patients and 549 ethnically matched healthy controls were included in the initial study. A second sample set consisting of 294 CD patients and 475 controls was used for additional analysis. All these samples correspond to unrelated Spanish white individuals. CD patients were diagnosed following the European Society for Paediatric Gastroenterology and Nutrition (ESPGAN), 97% are positive for HLA-DQ2 and/or HLA-DQ8. Controls correspond mainly to blood donors and laboratory staff. CD samples were consecutively collected in two centres of the same region (Hospital La Paz and Hospital Clínico San Carlos, Madrid) and controls were collected at the Hospital Clínico San Carlos.

### Markers and genotyping

We selected genes with a known functional role in the Th17 immune response. Accordingly, sixteen genes were studied: *IL23R*, *RORC*, *IL6R*, *IL17A*, *IL17F*, *CCR6*, *IL6*, *JAK2*, *TNFSF15*, *IL23A*, *IL22*, *STAT3*, *TBX21*, *SOCS3*, *IL12RB1* and *IL17RA*. For all these genes except *IL6R*, *JAK2* and *STAT3*, single nucleotide polymorphisms (SNPs) were selected following the "aggressive tagging" option present in the Haploview program with genetic data downloaded from the HapMap Project (http://hapmap.ncbi.nlm.nih.gov) (50 kb upstream and downstream of the transcription initiation site). To increase statistical power, only markers with a minor allele frequency (MAF)>10% were included. In addition, SNPs located in those genes which code nonsynonymous changes or were previously associated with some autoimmune disease were also analysed independently of their MAF. In *STAT3* and *JAK2*, only two SNPs previously associated with Crohn's disease, which share some susceptibility factors with CD, were included: rs744166 and rs10758669, respectively; and in *IL6R* we studied one functional polymorphism, rs8192284. SNPs located in *IL6* and *IL6R* and two SNPs in *IL23R*, rs11209026 and rs7517847, were analysed in previous works [12,23], which included most of the samples initially analysed in this study, but their data were used to evaluate genetic interactions with other Th17 related genes.

A total of 101 SNPs were initially studied (Table S1). All of them were genotyped by Veracode technology performed at the National Genotyping Center (http://www.cegen.org), except those that failed (rs10494269, rs9395767, rs608137, rs6927645, rs273506 and rs2241044) and those located in the *IL6*, *IL6R* and *TBX21* genes, which were genotyped with specific TaqMan assays. Two SNPs (rs11209026 and rs7517847, both in the *IL23R* gene) were genotyped by those two technologies and identical results were obtained.

Additional analysis included the study by TaqMan technology of rs12070470, in the *IL23R* gene.

### Statistical analysis

Deviations from Hardy-Weinberg proportions were assessed in all the SNPs studied.

A case-control analysis using chi-square tests was performed for each individual SNP and for the haplotypes resulting from the SNPs studied in the same genetic region.

Interactions between genes were evaluated following four different approaches: logistic regression, random forests (RF), classification and regression trees (CART) and multifactor dimensionality reduction (MDR).

## Results

Three SNPs showed deviation from Hardy-Weinberg proportions and were eliminated from the study: rs2064331 (*IL17F*), rs10878804 (*IL22*) and rs9645406 (*RORC*).

The comparison of genotypic frequencies between cases and controls for all the SNPs analysed achieved a nominal significant value in twelve polymorphisms located in eight different genetic regions (Table 1). Although none of them withstand Bonferroni correction, we tried to replicate associations involving *SOCS3* and *IL23R* using a second sample set. These two genes show the lowest case-control p-values in the present analysis and additionally some SNP in those genes showed a nominal significance in previous CD GWAS [20].

The initial *IL23R* data analysis also evidenced one haplotype significantly associated with CD susceptibility (rs4655683-rs10889667-rs1569922-rs790632-rs7517847-rs10489629-rs7528924-rs2201841-rs4655530-rs11209026-rs6682033-rs6693831, G-C-C-C-C-T-A-G-T-A-G-A-C): 9.2% in CD patients vs. 6.3% in controls ($p = 0.0067$). For replication purposes, the SNP rs12070470, highly correlated with that haplotype ($r^2 = 1$ according to http://hapmap.ncbi.nlm.nih.gov/) was studied in the second sample set instead of the 12 SNPs initially considered.

**Table 1.** Genetic polymorphisms which showed a nominal significant value after case-control comparisons (in decreasing significance).

| GENE | SNP | GENOTYPE | p | OR | 95% CI |
|---|---|---|---|---|---|
| *SOCS3* | rs4969170 | AA | 0.0018 | 0.59 | 0.42–0.84 |
| *IL23R* | rs7528924 | GG | 0.0057 | 2.11 | 1.19–3.74 |
| *TNFSF15* | rs17219926 | CC | 0.0103 | 1.43 | 1.08–1.89 |
| *IL6* | rs2069827 | GT+TT | 0.016 | 1.51 | 1.06–2.14 |
| *IL22* | rs11611206 | AA | 0.019 | 0.39 | 0.16–0.93 |
| *IL23R* | rs11209026 | AG+GG | 0.026 | 1.42 | 1.03–1.97 |
| *RORC* | rs1521186 | AA+AG | 0.027 | 1.31 | 1.02–1.67 |
| *IL22* | rs11177131 | CT+TT | 0.034 | 0.76 | 0.58–0.99 |
| *IL17A* | rs8193036 | CT+TT | 0.034 | 0.60 | 0.36–0.99 |
| *IL6* | rs1800795 | CG+CC | 0.037 | 1.26 | 1.01–1.58 |
| *TNFSF15* | rs6478108 | CT+CC | 0.043 | 0.79 | 0.63–1.00 |
| *CCR6* | rs3798315 | TT | 0.044 | 4.17 | 0.90–38.84 |

ORs are referred to the mutant genotype or carrier of the mutant allele (specified below "genotype").

doi:10.1371/journal.pone.0031244.t001

**Table 2.** Genotypic data (N (%)) for rs4969170 in the original and the replication sets.

| | Original set* | | Replication set# | |
|---|---|---|---|---|
| | CD | Controls | CD | Controls |
| | (N = 732) | (N = 551) | (N = 294) | (N = 462) |
| GG | 292 (39.9) | 212 (38.5) | 124 (42.2) | 198 (42.9) |
| AG | 368 (50.3) | 253 (45.9) | 138 (46.9) | 199 (43.1) |
| AA | 72 (9.8) | 86 (15.6) | 32 (10.9) | 65 (14.1) |

AA genotype:
*: p = 0.0018 OR = 0.59 95% CI 0.42–0.84;
#, p = 0.20 95% CI OR = 0.75 (0.46–1.20).
doi:10.1371/journal.pone.0031244.t002

No significant associations involving *IL23R* were observed in the replication set. Regarding the SNP rs4969170, in the *SOCS3* gene, a significant association was observed pooling the original and the replication sets: p = 0.0012 OR = 0.64 95% CI 0.49–0.84 (Table 2). Statistical power limitations probably precluded us to obtain a significant result in the replication set.

No consensus exists as to the best methodology to evaluate epistasis. Therefore, we used four different statistical methods to evaluate genetic interactions between all the studied polymorphisms located in different genes. We did not find statistically significant results with any methodological approach.

## Discussion

With the development of genome wide association studies (GWAS), the number of discovered genes involved in CD susceptibility has highly increased. However, the percentage of disease heritability explained has not experienced such an increase. Genetic variants not included in GWAS and genetic interactions could be underlying some missing heritability. We bear this in mind when studying the relevance of the Th17 immune response on CD susceptibility. We performed an extensive case-control study including sixteen genes which code relevant factors involved in that immune response. Tag SNPs were selected to cover most of the variability present in each gene, with exception of *IL6R*, *STAT3* and *JAK2*. SNPs coding nonsynonymous changes or those previously associated with other autoimmune diseases were also included in order to increase the *a priori* probability of obtaining a significant result. Additionally, we evaluated the possibility that interactions between the studied genes were involved in disease susceptibility. Our results seem to discard a relevant role of Th17 cells on CD risk, since no significantly associated SNP or gene-gene interaction was consistently observed, with the only exception of rs4969170, located in *SOCS3*, which deserves further research. However, although *SOCS3* is later confirmed, its functional role must be elucidated, since it is involved in different functional pathways and it would be expected that more than one Th17 gene was associated with CD susceptibility, as it has been observed with other Th17-mediated diseases.

The discovery of the IL-23 cytokine prompted the re-examination of the dominant Th response in many autoimmune diseases, primarily in those considered as skewed towards a Th1 phenotype. Studies based on murine models of multiple sclerosis, rheumatoid arthritis and inflammatory bowel disease related these conditions with a Th17 response. However, a similar conclusion is not drawn from GWAS results [24,25,26]. Although several Th17-related genes have been associated with IBD and RA, the list of MS susceptibility genes does not suggest a Th17 related etiology. This intriguing issue is probably far away from being answered. Nowadays it seems clear that Th immune responses are not independent and plasticity exists between Th cell subsets. A shift between Th1 and Th17 can occur during the inflammatory process and it is possible to speculate that the relative contribution and the timing of each subset will determine which genes would be involved in disease risk. Moreover, the cytokine microenvironment can determine the shift towards a specific immune response. From this point of view, genetics could not be so relevant if other compensatory mechanisms exist. This evidence, as previously suggested, that overlap between autoimmune diseases must be observed with caution. Th17 cells seem to mediate several autoimmune diseases but their impact in disease etiology seems to be different.

In summary, gene expression studies link CD pathogenesis to Th17 cells, but we evidenced that polymorphisms in Th17-related genes do not seem to be crucial for disease development. This is concordant with observations on MS. Although, in general, genetic data provide clues that ratified by functional studies unravel disease pathogenesis, this time it makes necessary to do somehow the other way around, with the special difficulty of explaining the divergent genetic results observed in different immune mediated diseases. Therefore, much more work is expected in this field.

## Supporting Information

**Table S1** Genes and SNPs studied ordered by chromosome and position.
(DOC)

## Acknowledgments

We are most grateful to Carmen Martínez Cuervo and M. Ángel García Martínez for their expert technical assistance.

## Author Contributions

Conceived and designed the experiments: CN. Performed the experiments: LMM BD LE. Analyzed the data: LMM MGM BD. Contributed reagents/materials/analysis tools: MGM CM IP MFA. Wrote the paper: CN LMM CM. Critically revised the manuscript: LMM MGM BD LE CM IP MAF CN.

## References

1. Oukka M (2008) Th17 cells in immunity and autoimmunity. Ann Rheum Dis 67 Suppl 3: iii26–29.
2. Volpe E, Servant N, Zollinger R, Bogiatzi SI, Hupe P, et al. (2008) A critical function for transforming growth factor-beta, interleukin 23 and proinflammatory cytokines in driving and modulating human T(H)-17 responses. Nat Immunol 9: 650–657.
3. Cua DJ, Sherlock J, Chen Y, Murphy CA, Joyce B, et al. (2003) Interleukin-23 rather than interleukin-12 is the critical cytokine for autoimmune inflammation of the brain. Nature 421: 744–748.
4. Yen D, Cheung J, Scheerens H, Poulet F, McClanahan T, et al. (2006) IL-23 is essential for T cell-mediated colitis and promotes inflammation via IL-17 and IL-6. J Clin Invest 116: 1310–1316.
5. Duerr RH, Taylor KD, Brant SR, Rioux JD, Silverberg MS, et al. (2006) A genome-wide association study identifies IL23R as an inflammatory bowel disease gene. Science 314: 1461–1463.
6. Burton PR, Clayton DG, Cardon LR, Craddock N, Deloukas P, et al. (2007) Association scan of 14,500 nonsynonymous SNPs in four diseases identifies autoimmunity variants. Nat Genet 39: 1329–1337.

7. Cargill M, Schrodi SJ, Chang M, Garcia VE, Brandon R, et al. (2007) A large-scale genetic association study confirms IL12B and leads to the identification of IL23R as psoriasis-risk genes. Am J Hum Genet 80: 273–290.
8. Steinman L (2007) A brief history of T(H)17, the first major revision in the T(H)1/T(H)2 hypothesis of T cell-mediated tissue damage. Nat Med 13: 139–145.
9. Lees CW, Barrett JC, Parkes M, Satsangi J (2011) New IBD genetics: common pathways with other diseases. Gut Epub ahead of print.
10. Glas J, Stallhofer J, Ripke S, Wetzke M, Pfennig S, et al. (2009) Novel genetic risk markers for ulcerative colitis in the IL2/IL21 region are in epistasis with IL23R and suggest a common genetic background for ulcerative colitis and celiac disease. Am J Gastroenterol 104: 1737–1744.
11. McGovern DP, Rotter JI, Mei L, Haritunians T, Landers C, et al. (2009) Genetic epistasis of IL23/IL17 pathway genes in Crohn's disease. Inflamm Bowel Dis 15: 883–889.
12. Nunez C, Dema B, Cenit MC, Polanco I, Maluenda C, et al. (2008) IL23R: a susceptibility locus for celiac disease and multiple sclerosis? Genes Immun 9: 289–293.
13. Harris KM, Fasano A, Mann DL (2008) Cutting edge: IL-1 controls the IL-23 response induced by gliadin, the etiologic agent in celiac disease. J Immunol 181: 4457–4460.
14. Einarsdottir E, Koskinen LL, Dukes E, Kainu K, Suomela S, et al. (2009) IL23R in the Swedish, Finnish, Hungarian and Italian populations: association with IBD and psoriasis, and linkage to celiac disease. BMC Med Genet 10: 8.
15. Monteleone I, Sarra M, Del Vecchio Blanco G, Paoluzi OA, Franze E, et al. (2011) Characterization of IL-17A-producing cells in celiac disease mucosa. J Immunol 184: 2211–2218.
16. Castellanos-Rubio A, Santin I, Irastorza I, Castano L, Carlos Vitoria J, et al. (2009) TH17 (and TH1) signatures of intestinal biopsies of CD patients in response to gliadin. Autoimmunity 42: 69–73.
17. Fernandez S, Molina IJ, Romero P, Gonzalez R, Pena J, et al. (2011) Characterization of gliadin-specific Th17 cells from the mucosa of celiac disease patients. Am J Gastroenterol 106: 528–538.
18. van Heel DA, Franke L, Hunt KA, Gwilliam R, Zhernakova A, et al. (2007) A genome-wide association study for celiac disease identifies risk variants in the region harboring IL2 and IL21. Nat Genet 39: 827–829.
19. Hunt KA, Zhernakova A, Turner G, Heap GA, Franke L, et al. (2008) Newly identified genetic risk variants for celiac disease related to the immune response. Nat Genet 40: 395–402.
20. Dubois PC, Trynka G, Franke L, Hunt KA, Romanos J, et al. (2010) Multiple common variants for celiac disease influencing immune gene expression. Nat Genet 42: 295–302.
21. Bush WS, McCauley JL, DeJager PL, Dudek SM, Hafler DA, et al. (2011) A knowledge-driven interaction analysis reveals potential neurodegenerative mechanism of multiple sclerosis susceptibility. Genes Immun 12: 335–340.
22. Torok HP, Glas J, Endres I, Tonenchi L, Teshome MY, et al. (2009) Epistasis between Toll-like receptor-9 polymorphisms and variants in NOD2 and IL23R modulates susceptibility to Crohn's disease. Am J Gastroenterol 104: 1723–1733.
23. Dema B, Martinez A, Fernandez-Arquero M, Maluenda C, Polanco I, et al. (2009) The IL6-174G/C polymorphism is associated with celiac disease susceptibility in girls. Hum Immunol 70: 191–194.
24. Franke A, McGovern DP, Barrett JC, Wang K, Radford-Smith GL, et al. (2010) Genome-wide meta-analysis increases to 71 the number of confirmed Crohn's disease susceptibility loci. Nat Genet 42: 1118–1125.
25. Stahl EA, Raychaudhuri S, Remmers EF, Xie G, Eyre S, et al. (2010) Genome-wide association study meta-analysis identifies seven new rheumatoid arthritis risk loci. Nat Genet 42: 508–514.
26. Sawcer S, Hellenthal G, Pirinen M, Spencer CC, Patsopoulos NA, et al. (2011) Genetic risk and a primary role for cell-mediated immune mechanisms in multiple sclerosis. Nature 476: 214–219.

Human Immunology xxx (2013) xxx–xxx

Contents lists available at ScienceDirect

journal homepage: www.elsevier.com/locate/humimm

# Role of *TNFRSF1B* polymorphisms in the response of Crohn's disease patients to infliximab

L.M. Medrano [a], C. Taxonera [b], A. Márquez [c], M. Barreiro-de Acosta [d], M. Gómez-García [e], C. González-Artacho [e], J.L. Pérez-Calle [f], F. Bermejo [g], A. Lopez-Sanromán [h], M.D. Martín Arranz [i], J.P. Gisbert [j], J.L. Mendoza [b], J. Martín [c], E. Urcelay [a], C. Núñez [a,*]

[a] *Immunology Department, Hospital Clínico San Carlos, Instituto de Investigación Sanitaria del Hospital Clínico San Carlos (IdISSC), Madrid, Spain*
[b] *Gastroenterology Department, Hospital Clínico San Carlos, Instituto de Investigación Sanitaria del Hospital Clínico San Carlos (IdISSC), Madrid, Spain*
[c] *Institute of Parasitology and Biomedicine, CSIC, Granada, Spain*
[d] *Department of Gastroenterology, Hospital Clínico Universitario de Santiago de Compostela, Spain*
[e] *Department of Gastroenterology, Hospital Virgen de las Nieves, Granada, Spain*
[f] *Department of Gastroenterology, Hospital Alcorcón, Madrid, Spain*
[g] *Department of Gastroenterology, Hospital Fuenlabrada, Madrid, Spain*
[h] *Department of Gastroenterology, Hospital Ramón y Cajal, Madrid, Spain*
[i] *Department of Gastroenterology, Hospital La Paz, Madrid, Spain*
[j] *Department of Gastroenterology, Hospital Universitario de La Princesa, Instituto de Investigación Sanitaria Princesa (IP), Centro de Investigación Biomédica en Red de Enfermedades Hepáticas y Digestivas (CIBERHED), Madrid, Spain*

ARTICLE INFO

*Article history:*
Received 1 July 2013
Accepted 27 September 2013
Available online xxxx

ABSTRACT

Infliximab (IFX) is a valid treatment for Crohn's disease (CD), but a relevant percentage of patients do not benefit from this therapy. In the Japanese population, the response to IFX was associated with markers in the TNF receptor superfamily 1A (*TNFRSF1A*) and 1B (*TNFRSF1B*) genes. We aimed to replicate the association previously described in the Japanese population and to ascertain the role of TNF receptors as modulators of the response to IFX. We studied 297 white Spanish CD patients with a known response to IFX: 238 responders and 59 primary nonresponders. Four single nucleotide polymorphisms (SNPs) were analyzed: rs767455 in *TNFRSF1A* and rs1061622, rs1061624, and rs3397 in *TNFRSF1B*. Comparisons between groups were performed with chi-square tests or the Fisher's exact test. Different features (sex, age, disease duration, smoking among others) were evaluated as possible confounding factors. No significant association was found between the studied *TNFRSF1A* polymorphisms and response to IFX. In the *TNFRSF1B* gene, the haplotype rs1061624_A-rs3397_T was significantly increased in nonresponders: $p = 0.015$, OR = 1.78, 95% CI 1.09–2.90; and an increased frequency of rs1061622_G carriers was observed in patients with remission: $p = 0.033$ vs nonresponders and $p = 0.023$ vs patients with a partial response. Our results support a role of *TNFRSF1B* gene variants in the response to IFX in CD patients.


## 1. Introduction

Crohn's disease (CD) is characterized by a chronic inflammation of the gastrointestinal tract, which is most probably caused by an inappropriate immunological response to intestinal microbes in genetically susceptible individuals [1]. Recently, many genes have been associated with risk of disease, mainly due to the development of genome wide association studies (GWAS) [2]. Genetic factors also underlie the observed differences among individuals in response to the drugs used as therapy. This opens up the possibility of searching for genetic markers to assist us in the classification of patients as responders or nonresponders to specific treatments. Recent advances in CD treatments make such a classification particularly important. Ideally, the best treatment for each patient could be determined or, at least, the individual chance of success for a specific treatment could be predicted.

Tumour necrosis factor alpha (TNF-α) is a nonspecific mediator of tissue injury in CD. It can bind to two different receptors: *TNFRSF1A* (tumour necrosis factor receptor superfamily 1A) and *TNFRSF1B* (tumour necrosis factor receptor superfamily 1B); which activates NF-κB and triggers several inflammatory pathways [3]. Infliximab (IFX) is a chimeric monoclonal IgG1 that neutralizes TNF by binding it [4]. IFX has been proven to be effective for the

* Corresponding author. Address: Immunology Department, Hospital Clínico San Carlos, Instituto de Investigación Sanitaria del H. Clínico San Carlos (IdISSC). Profesor Martín Lagos s/n, 28040 Madrid, Spain. Fax: +34 91 330 3344.
*E-mail addresses:* conchita.npardo@gmail.com, mariaconcepcion.nunez@salud.madrid.org (C. Núñez).

0198-8859/$36.00 - see front matter © 2013 American Society for Histocompatibility and Immunogenetics. Published by Elsevier Inc. All rights reserved.
http://dx.doi.org/10.1016/j.humimm.2013.09.017

treatment of both luminal [5] and fistulizing CD [6]. However, around 20–30% of CD patients do not show a good response to IFX [7]. In addition to the genetic factors underlying this variable response, other parameters such as previous surgery, concomitant treatments, smoking habits or disease duration may contribute to the observed variability.

The widespread use of IFX for CD treatment has led to an increase in the number of pharmacogenetic studies. The binding affinity and/or the kinetic stability of TNF for their receptors *TNFRSF1A* or *TNFRSF1B* could be affected by the presence of genetic variants in these receptors. In addition, polymorphisms in the genes coding these receptors could affect the signal transduction cascade. Therefore, TNF receptors are good candidates as response modifiers. Scarce information exist about the functional role of genetic polymorphisms in *TNFRSF1A* and *TNFRSF1B*, but it seems that the non-synonymous polymorphism rs1061622 and some haplotypes conformed by polymorphisms located in the 3′-untranslated region (UTR) (rs1061624, rs5030792, and rs3397), all related to the *TNFRSF1B* gene, may have functional consequences through their role in signal transduction or in mRNA stability [8–10]. The role of some genetic polymorphisms in *TNFRSF1A*, located on chromosome 12p13 and *TNFRSF1B*, located on 1p36, in the response of CD patients to IFX have been previously studied, but consensus is lacking. In the Japanese population, genetic variants in *TNFRSF1A* and *TNFRSF1B*, including the previously described ones with a functional role, were studied and association with primary nonresponse to IFX was found in a sample of 80 patients, 32 of them classified as nonresponders using the Harvey–Bradshaw index (HBI) [11]. In Caucasian populations, apparently discrepant results have been reported [12–14] in studies with a variable sample size, but different criteria to evaluate response to IFX were followed (Crohn's disease activity index, CDAI; overall clinical evaluation). Most of these studies are only focused on primary nonresponse to IFX. A secondary loss of response is also relevant, but it probably has a different etiology, as it is caused by different mechanisms, mainly pharmacodynamics mechanisms but also by immunogenicity and by non-immune-mediated pharmacokinetics [15].

We aimed to study the genetic polymorphisms in *TNFRSF1A* and *TNFRSF1B* previously associated with response to IFX in the Japanese population in our sample of IFX treated CD patients.

## 2. Materials and methods

### 2.1. Patients and criteria of response to IFX

We collected samples from CD patients treated with IFX from 7 centers around Spain: 238 responders and 59 primary nonresponders to this therapy. All of them were Spanish of Caucasian ancestry and at least 18 years old of age. All had been diagnosed with CD according to clinical, radiologic, endoscopic, and histological criteria [16] and had received at least three induction doses of IFX (5 mg per kilogram) at weeks 0, 2, and 6. IFX was administered to treat either moderate to severe active luminal CD or active fistulizing perianal CD. All patients gave written informed consent and the study was approved by the ethic committees of all participating hospitals.

The response to IFX was determined by a chronological review of the medical records, and data were centrally monitored. In patients with luminal disease, response was evaluated by the Harvey–Bradshaw index (HBI) [17] at the beginning and 10 weeks after the first IFX dose. Partial response was defined as a decrease in the HBI of more than 3 points and absence of concomitant corticosteroids. Remission was defined as a final HBI < 4 and absence of concomitant corticosteroids. In patients with perianal disease, response was evaluated at week 10 after the first IFX dose. Remission was defined as the complete closure of all fistulas and partial response as a reduction ($\geqslant$50%) in the number of draining fistulas. Patients who received IFX for both luminal and fistulizing disease and who achieved remission of any type (in the luminal or in the fistulizing disease) that justified maintenance treatment with IFX were considered as responders. All patients who did not achieve partial response or remission after the three IFX induction doses were considered as primary nonresponders.

### 2.2. Genotyping

We analyzed four SNPs previously studied in the Japanese population within the *TNFRSF1A* and *TNFRSF1B* genes: rs767455 in a coding exon of *TNFRSF1A*; and rs1061622, rs1061624, and rs3397 in *TNFRSF1B*. Genotyping was performed by TaqMan technology (Applied Biosystems, Foster City, CA, USA) and at least 95% of samples were successfully genotyped for each SNP. Haplotypes were estimated using the EM (Expectation–Maximization) algorithm from the Haploview 4.1 software.

### 2.3. Statistical analyses

Genetic comparisons between responders and primary nonresponders were analyzed by chi-square tests or the Fisher's exact test (when the expected values were lower than five). Since we aimed to replicate previously described associations, no correction for multiple testing was necessary.

Comparisons between CD groups for other clinical or demographic characteristics corresponding to first IFX infusion (Table 1) were performed using chi-square tests for categorical variables and the Mann–Whitney *U* test for continuous variables. Those variables (sex, disease duration, smoking at first administration dose, concomitant immunomodulator treatment, etc.) in addition to partial or total response to IFX were evaluated as possible confounding factors by logistic regression using the SPSS 15.0 software. Concomitant immunomodulator treatment was defined as the use of any immunosuppressive agent (azathioprine, mercaptopurine, methotrexate, and mycophenolate mofetil) for at least 4 weeks before the first infliximab induction dose.

## 3. Results

All the polymorphisms included in the study conformed to Hardy–Weinberg expectations. Linkage disequilibrium between the SNPs in *TNFRSF1B* was as follows: $D' = 0.25$ $r^2 = 0.012$ LOD = 0.64 between rs1061622 and rs1061624; $D' = 0.10$ $r^2 = 0.001$ LOD = 0.08 between rs1061622 and rs3397; and $D' = 0.40$ $r^2 = 0.114$ LOD = 7.8 between rs1061624 and rs3397. Therefore, only haplotypes conformed by rs1061624 and rs3397, the two SNPs showing significant linkage disequilibrium, were considered.

### 3.1. Base-line characteristics

The comparison of several demographic and clinical characteristics between responders and nonresponders to IFX showed a significant difference for disease duration (Table 1).

### 3.2. Response vs nonresponse to IFX

The genetic frequencies for all the SNPs in responders and nonresponders are shown in Table 2. We did not observe significant differences between responders and nonresponders when studying the SNP located in the *TNFRSF1A* gene. In *TNFRSF1B*, the frequency of the allele A of rs1061624 was significantly higher in

Please cite this article in press as: Medrano LM et al. Role of *TNFRSF1B* polymorphisms in the response of Crohn's disease patients to infliximab. Hum Immunol (2013), http://dx.doi.org/10.1016/j.humimm.2013.09.017

**Table 1**
Characteristics of Crohn's disease patients at first infliximab infusion, both responders ($n = 238$) and nonresponders ($n = 59$) to infliximab.

| | Nonresponders | | Responders | | *p* value |
|---|---|---|---|---|---|
| | *N* | % | *N* | % | |
| *Age* | 43.1 ± 1.6 | | 39.9 ± 0.8 | | 0.619 |
| *Sex* | | | | | |
| Male | 26 | 41.0 | 109 | 41.1 | 0.81 |
| Female | 33 | 59.0 | 129 | 58.9 | |
| *Disease duration (years)* | 13.5 ± 1.1 | | 10.6 ± 0.5 | | 0.007 |
| *Age at diagnosis (A)* | | | | | |
| A1 | 6 | 9.3 | 31 | 13.7 | |
| A2 | 45 | 79.1 | 175 | 72.6 | 0.83 |
| A3 | 7 | 11.6 | 29 | 13.7 | |
| *Location (L)* | | | | | |
| L1 | 23 | 40.5 | 64 | 25.6 | 0.174[a] |
| L2 | 9 | 14.3 | 42 | 16.9 | |
| L3 | 25 | 45.0 | 117 | 55.6 | |
| L4 | 0 | 0 | 0 | 0 | |
| L1 + L4 | 0 | 0 | 3 | 1.3 | |
| L2 + L4 | 0 | 0 | 0 | 0 | |
| L3 + L4 | 0 | 0 | 5 | 0.62 | |
| *Behavior (B)* | | | | | |
| B1 | 18 | 28.6 | 66 | 23.0 | 0.764 |
| B2 | 5 | 11.9 | 17 | 10.0 | |
| B3 | 9 | 16.7 | 33 | 12.5 | |
| B1p | 7 | 14.3 | 59 | 26.9 | |
| B2p | 2 | 0 | 10 | 5.0 | |
| B3p | 15 | 28.6 | 47 | 22.5 | |
| *Smoking* | | | | | |
| Yes | 20 | 31.8 | 105 | 47.7 | 0.076 |
| No | 39 | 68.2 | 116 | 52.3 | |
| *Immunomodulator treatment*[b] | | | | | |
| Yes | 44 | 93.5 | 172 | 92.9 | 0.464 |
| No | 3 | 6.5 | 16 | 7.1 | |

A1: ⩽ 16 years; A2: 17–40 years; A3: >40 years.
L1: terminal ileum; L2: colon; L3: ileocolon; L4: Upper GI; L1 + L4: terminal ileum + upper GI; L2 + L4: colon + upper GI; L3 + L4: ileocolon + upper GI.
B1: nonstricturing, nonpenetrating; B2: structuring; B3: penetrating; B1p: nonstricturing, nonpenetrating + perianal; B2p: stricturing + perianal; B3p: penetrating + perianal.
[a] Excluding categories with L4.
[b] Patients with no concomitant treatment have not been included in this comparison.

nonresponders (53.6%) than in responders to IFX (41.5%) ($p = 0.02$; OR = 1.63 95% CI 1.05–2.51). When studying rs3397, a statistically significant result was observed for the minor genotype (CC), which was increased in responders to IFX: $p = 0.05$, OR = 3.19 95% CI 0.95–16.78. The analysis of the haplotypes conformed by these two SNPs showed one haplotype, rs1061624_A rs3397_T, significantly increased in nonresponders ($p = 0.015$, OR = 1.78 95% CI 1.09–2.90) (Table 4).

### 3.3. Partial response to IFX vs remission

In Table 3, genotypic and allelic frequencies of the studied SNPs in patients showing partial response or remission after treatment with IFX are presented for the available patients. Significant differences are only observed when studying rs1061622. Comparison of each of those groups with nonresponders showed only significant differences when considering patients with remission: 29% carriers of the G allele in nonresponders vs 47% in patients showing remission ($p = 0.033$) and vs 23% in patients with a partial response ($p = 0.55$).

All the factors listed in Table 1 are tested as potential confounders of the association between the individual SNPs and the haplotypes and response to IFX, but none of them showed any effect.

**Table 2**
Genotypic and allelic frequencies for the studied SNPs.

| | Nonresponders | | Responders | | *p* Value | OR (95% CI) |
|---|---|---|---|---|---|---|
| | *N* | % | *N* | % | | |
| *TNFRSF1A* | | | | | | |
| rs767455 | | | | | | |
| AA | 25 | 45.5 | 99 | 43.4 | – | 1.0 (reference) |
| AG | 23 | 41.8 | 94 | 41.2 | 0.92 | 0.97 (0.49–1.91) |
| GG | 7 | 12.7 | 35 | 15.4 | 0.62 | 0.79 (0.28–2.15) |
| AG + GG | 30 | 54.5 | 129 | 56.6 | 0.78 | 0.92 (0.49–1.74) |
| A | 73 | 66.4 | 292 | 64.0 | 0.65 | 0.90 (0.57–1.43) |
| G | 37 | 33.6 | 164 | 36.0 | | |
| *TNFRSF1B* | | | | | | |
| rs1061622 | | | | | | |
| TT | 41 | 70.7 | 144 | 61.3 | – | 1.0 (reference) |
| TG | 16 | 27.6 | 84 | 35.7 | 0.21 | 0.67 (0.34–1.32) |
| GG | 1 | 1.7 | 7 | 3.0 | 0.51 | 0.50 (0.01–4.10) |
| TG + GG | 17 | 29.3 | 91 | 38.7 | 0.18 | 0.66 (0.34–1.27) |
| T | 98 | 84.5 | 372 | 79.1 | 0.20 | 0.70 (0.39–1.24) |
| G | 18 | 17.5 | 98 | 20.9 | | |
| rs1061624 | | | | | | |
| GG | 10 | 17.9 | 88 | 37.5 | – | 1.0 (reference) |
| GA | 32 | 57.1 | 99 | 42.1 | 0.006 | 2.84 (1.25–6.59) |
| AA | 14 | 25.0 | 48 | 20.4 | 0.033 | 2.57 (0.98–6.80) |
| GA + AA | 46 | 82.1 | 147 | 62.5 | 0.005 | 2.75 (1.26–6.15) |
| G | 52 | 46.4 | 275 | 58.5 | 0.021 | 1.63 (1.05–2.51) |
| A | 60 | 53.6 | 195 | 41.5 | | |
| rs3397 | | | | | | |
| CC | 3 | 5.2 | 35 | 14.8 | – | 1.0 (reference) |
| TC | 30 | 51.7 | 101 | 42.8 | 0.040 | 3.47 (0.98–18.73) |
| TT | 25 | 43.1 | 100 | 42.4 | 0.083 | 2.92 (0.81–15.93) |
| TT + TC | 55 | 94.8 | 201 | 85.2 | 0.049 | 3.19 (0.95–16.78) |
| T | 80 | 69.0 | 301 | 63.8 | 0.29 | 0.79 (0.50–1.25) |
| C | 36 | 31.0 | 171 | 36.2 | | |

The genotype indicated as OR = 1.0 is used as the reference for the remaining genotypes and for the carrier to calculate *p* values and OR.
Allelic ORs are relative to the minor allele in responders.

## 4. Discussion

We studied the association of several genetic polymorphisms located in *TNFRSF1A* and *TNFRSF1B* with response to IFX in CD patients. When considering several demographic and clinical characteristics, a better response to IFX was observed in patients with an earlier treatment, as previously described [18]. In addition, we found that one haplotype conformed by two SNPs studied in *TNFRSF1B* (rs1061624_A-rs3397_T) was significantly increased in nonresponders to this therapy. A similar result was previously observed in the Japanese population following the same criteria of response to IFX that one used in this work [11], which suggest the existence of a genetic variant in linkage disequilibrium with this haplotype modulating the response to IFX. We also found a genetic variant, rs1061622, significantly increased in patients showing remission after IFX treatment. Steenholdt et al. [14] also reported an increased frequency of carriage of rs1061622_C in this particular group of patients. It is very interesting that a functional role has been described for one haplotype containing the rs1061624 and rs3397 polymorphisms [9], and for the variant rs1061622.

The role of genetic variants in the genes coding for TNF receptors as predictors of response to IFX for CD treatment has been previously studied by several groups [11–14]. It should be highlighted that all those studies included a reduced number of SNPs and therefore no conclusive results could be obtained. However, lack of association between *TNFRSF1A* polymorphisms and response to IFX has been quite consistently found, which dampened interest in this gene. The case of *TNFRSF1B* is different, with several polymorphisms associated with response to IFX by different groups. The haplotype rs1061624_A-rs3397_T was initially associated with

Please cite this article in press as: Medrano LM et al. Role of *TNFRSF1B* polymorphisms in the response of Crohn's disease patients to infliximab. Hum Immunol (2013), http://dx.doi.org/10.1016/j.humimm.2013.09.017

**Table 3**
Genotypic and allelic frequencies for the studied SNPs in patients stratified by type of response.

| | Partial response | | Remission | | *p* Value | OR (95%CI) |
|---|---|---|---|---|---|---|
| | *N* | % | *N* | % | | |
| *TNFRSF1A* | | | | | | |
| rs767455 | | | | | | |
| AA | 14 | 45.2 | 40 | 40.8 | – | 1.0 (reference) |
| AG | 13 | 41.9 | 39 | 39.8 | 0.91 | 0.95 (0.36–2.49) |
| GG | 4 | 12.9 | 19 | 19.4 | 0.42 | 0.60 (0.13–2.29)e-3.53) |
| AG + GG | 17 | 54.8 | 58 | 59.2 | 0.67 | 0.84 (0.34–2.04) |
| A | 41 | 66.1 | 119 | 60.7 | 0.44 | 0.79 (0.42–1.50) |
| G | 21 | 33.9 | 77 | 39.3 | | |
| *TNFRSF1B* | | | | | | |
| rs1061622 | | | | | | |
| TT | 23 | 76.7 | 54 | 53.5 | – | 1.0 (reference) |
| TG | 6 | 20.0 | 45 | 44.6 | 0.016 | 0.31 (0.10–0.89) |
| GG | 1 | 3.3 | 2 | 2.0 | 0.898 | 1.17 (0.02–23.56) |
| TG + GG | 7 | 23.3 | 47 | 46.6 | 0.023 | 0.35 (0.12–0.96) |
| T | 52 | 86.7 | 153 | 75.7 | 0.071 | 0.48 (0.20–1.14) |
| G | 8 | 13.3 | 49 | 24.3 | | |
| rs1061624 | | | | | | |
| GG | 11 | 35.5 | 37 | 36.3 | – | 1.0 (reference) |
| GA | 11 | 35.5 | 41 | 40.2 | 0.83 | 0.90 (0.32–2.57) |
| AA | 9 | 29.0 | 24 | 23.5 | 0.66 | 1.26 (0.40–3.92) |
| GA + AA | 19 | 64.5 | 65 | 63.7 | 0.97 | 0.98 (0.39–2.49) |
| G | 33 | 53.2 | 115 | 56.4 | 0.66 | 1.14 (0.62–2.09) |
| A | 29 | 46.8 | 89 | 43.6 | | |
| rs3397 | | | | | | |
| CC | 4 | 12.9 | 18 | 17.6 | – | 1.0 (reference) |
| TC | 11 | 35.5 | 44 | 43.1 | 0.86 | 1.13 (0.28–5.48) |
| TT | 16 | 51.6 | 40 | 39.2 | 0.34 | 1.80 (0.48–8.39) |
| TT + TC | 27 | 87.1 | 84 | 82.3 | 0.53 | 1.45 (0.42–6.37) |
| T | 43 | 69.4 | 124 | 60.8 | 0.22 | 0.68 (0.36–1.31) |
| C | 19 | 30.6 | 80 | 39.2 | | |

Partial response and remission data are shown for the available patients.
The genotype indicated as OR = 1.0 is used as the reference for the remaining genotypes and for the carrier to calculate *p* values and OR.
Allelic ORs are relative to the minor allele in responders.

**Table 4**
Frequency of the haplotypes conformed by the studied SNPS in *TNFRSF1B* in linkage disequilibrium (rs1061624, rs3397) in nonresponders ($2N = 118$) and in responders ($2N = 476$).

| Haplotype | Nonresponders (%) | Responders (%) | *p* value | OR (95%CI) |
|---|---|---|---|---|
| GT | 39.6 | 45.5 | 0.2496 | |
| AC | 24.6 | 23.2 | 0.7377 | |
| AT | 28.5 | 18.3 | 0.0142 | 1.78 (1.09–2.90) |
| GC | 7.3 | 13.1 | 0.0869 | |

IFX response in the Japanese population and no data exist in Caucasians. Mascheretti et al. [12] studied those two SNPs in the German population but haplotype analyses were not performed. These authors did not find an association after studying the individual SNPs; such association was not detected in the Japanese study either. It should be noted that the distribution of these alleles in their carrying haplotypes will influence the chance of obtaining a significant result, which will be obtained only when one allele is mainly present in the associated haplotype.

A high rate of false positive results has been suggested as responsible for the discrepant results observed in pharmacogenetic studies, which would be influenced by the inclusion of low numbers of patients. However, replication of a significant result in a different sample or population is warranted to validate a reported association and it avoids the necessity of correction for multiple testing. Thus, after our significant result, *TNFRSF1B* seems to contain a genetic variant involved in the clinical response to IFX. A role of this gene is also supported by the reported association of *TNFRSF1B* variants with biological response to IFX [13,14] and by the differences in *TNFRSF1B* expression observed depending on how the patients respond to this therapy [14,19]. These findings support the influence of this gene in the response of CD patients to IFX treatment.

It is also important to keep in mind that the apparent discrepancy between studies should be interpreted with caution because of differences in the criteria for evaluating response to IFX.

A limitation of the study could be a potential bias in the assessment of response due to the lack of endoscopy or MRI. In clinical trials the definition of response or remission has been based on clinical activity indexes and not on CRP values or endoscopic or MRI assessment [5,6].

The ultimate objective of pharmacogenetics, application of a personalized medicine, remains a distant goal given the low OR observed. Nevertheless, although the haplotype rs1061624_A-rs3397_T does not represent a useful pharmacogenetic marker, the situation might be different when the functional variant modulating the response to IFX is identified. Much more genetic variation than the one studied in *TNFRS1B* exists and a higher effect is expected for the functional variant. Unfortunately, the sample size of our study does not provide with enough statistical power to look for this variant. A large undertaking, most probably involving international collaboration, is now expected to attain this objective. The approach would involve collecting samples from a large number of CD patients with homogeneous criteria for IFX administration and for evaluation of the clinical response.

## Acknowledgments

We thank patients and controls for making this study feasible. Angel García and Carmen Martínez provided expert technical assistance.

## References

[1] Abraham C, Cho JH. Inflammatory bowel disease. N Engl J Med 2009;361:2066–78.
[2] Franke A, McGovern DP, Barrett JC, Wang K, Radford-Smith GL, Ahmad T, et al. Genome-wide meta-analysis increases to 71 the number of confirmed Crohn's disease susceptibility loci. Nat Genet 2010;42:1118–25.
[3] Monteleone G, Pallone F, MacDonald TT. Emerging immunological targets in inflammatory bowel disease. Curr Opin Pharmacol 2011;11:640–5.
[4] Van Assche G, Rutgeerts P. Anti-TNF agents in Crohn's disease. Expert Opin Investig Drugs 2000;9:103–11.
[5] Rutgeerts P, Feagan BG, Lichtenstein GR, Mayer LF, Schreiber S, Colombel JF, et al. Comparison of scheduled and episodic treatment strategies of infliximab in Crohn's disease. Gastroenterology 2004;126:402–13.
[6] Sands BE, Anderson FH, Bernstein CN, Chey WY, Feagan BG, Fedorak RN, et al. Infliximab maintenance therapy for fistulizing Crohn's disease. N Engl J Med 2004;350:876–85.
[7] Chaudhary R, Ghosh S. Prediction of response to infliximab in Crohn's disease. Dig Liver Dis 2005;37:559–63.
[8] Morita C, Horiuchi T, Tsukamoto H, Hatta N, Kikuchi Y, Arinobu Y, et al. Association of tumor necrosis factor receptor type II polymorphism 196R with systemic lupus erythematosus in the Japanese: molecular and functional analysis. Arthritis Rheum 2001;44:2819–27.
[9] Puga I, Lainez B, Fernandez-Real JM, Buxade M, Broch M, Vendrell J, et al. A polymorphism in the 3′ untranslated region of the gene for tumor necrosis factor receptor 2 modulates reporter gene expression. Endocrinology 2005;146:2210–20.
[10] Till A, Rosenstiel P, Krippner-Heidenreich A, Mascheretti-Croucher S, Croucher PJ, Schafer H, et al. The Met-196 -> arg variation of human tumor necrosis factor receptor 2 (TNFR2) affects TNF-alpha-induced apoptosis by impaired NF-kappaB signaling and target gene expression. J Biol Chem 2005;280:5994–6004.
[11] Matsukura H, Ikeda S, Yoshimura N, Takazoe M, Muramatsu M. Genetic polymorphisms of tumour necrosis factor receptor superfamily 1A and 1B affect responses to infliximab in Japanese patients with Crohn's disease. Aliment Pharmacol Ther 2008;27:765–70.
[12] Mascheretti S, Hampe J, Kuhbacher T, Herfarth H, Krawczak M, Folsch UR, et al. Pharmacogenetic investigation of the TNF/TNF-receptor system in patients

with chronic active Crohn's disease treated with infliximab. Pharmacogenomics J 2002;2:127–36.
[13] Pierik M, Vermeire S, Steen KV, Joossens S, Claessens G, Vlietinck R, et al. Tumour necrosis factor-alpha receptor 1 and 2 polymorphisms in inflammatory bowel disease and their association with response to infliximab. Aliment Pharmacol Ther 2004;20:303–10.
[14] Steenholdt C, Enevold C, Ainsworth MA, Brynskov J, Thomsen OO, Bendtzen K. Genetic polymorphisms of tumour necrosis factor receptor superfamily 1b and fas ligand are associated with clinical efficacy and/or acute severe infusion reactions to infliximab in Crohn's disease. Aliment Pharmacol Ther 2012;36:650–9.
[15] Steenholdt C, Brynskov J, Thomsen OO, Munck LK, Fallingborg J, Christensen LA, et al. Individualised therapy is more cost-effective than dose intensification in patients with Crohn's disease who lose response to anti-TNF treatment: a randomised, controlled trial. Gut 2013 [Epub ahead of print].
[16] Lennard-Jones JE. Classification of inflammatory bowel disease. Scand J Gastroenterol Suppl 1989;170:2–6 [discussion 16–19].
[17] Harvey RF, Bradshaw JM. A simple index of Crohn's-disease activity. Lancet 1980;1:514.
[18] Miheller P, Lakatos PL, Horvath G, Molnar T, Szamosi T, Czegledi Z, et al. Efficacy and safety of infliximab induction therapy in Crohn's disease in Central Europe – a Hungarian nationwide observational study. BMC Gastroenterol 2009;9:66.
[19] Arijs I, Quintens R, Van Lommel L, Van Steen K, De Hertogh G, Lemaire K, et al. Predictive value of epithelial gene expression profiles for response to infliximab in Crohn's disease. Inflamm Bowel Dis 2010;16:2090–8.

Please cite this article in press as: Medrano LM et al. Role of *TNFRSF1B* polymorphisms in the response of Crohn's disease patients to infliximab. Hum Immunol (2013), http://dx.doi.org/10.1016/j.humimm.2013.09.017